

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 14:24:01 ; Search time 16096.1 Seconds  
(without alignments)  
8776.084 Million cell updates/sec

Title: US-09-919-408A-1

Perfect score: 3453

Sequence: 1 GCGGCTGCTACCGGCGC.....AAAAAAAAAAAAAAAAAAAA 3453

Scoring table: IDENTITY NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 2889711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 577422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Result No.	Score	Query Match	Length	DB	ID	Description
1	3453	100.0	3453	6	AR005211	AR005211 Sequence
2	3453	100.0	3453	6	AR071703	AR071703 Sequence
3	3453	100.0	3453	6	125169	125169 Sequence 1
4	3453	100.0	3453	6	140600	140600 Sequence 1
5	3426	99.2	3426	6	AX695519	AX695519 Sequence
6	3426	99.2	3426	10	MUSFLK2	M64689 Mouse flk-2
7	3344	96.8	3521	6	144732	144732 Sequence 1
8	3344	96.8	3521	10	MMFLT3	X59398 Mouse Flt3
9	2979	86.3	2979	6	AX695520	AX695520 Sequence
10	2227.2	64.5	3501	6	AR005212	AR005212 Sequence
11	2227.2	64.5	3501	6	AR071704	AR071704 Sequence
12	2227.2	64.5	3501	6	125170	125170 Sequence 3
13	2227.2	64.5	3501	6	140601	140601 Sequence 3
14	2204.8	63.9	3476	6	140109	140109 Sequence 1
15	2203	63.8	3475	6	AX695522	AX695522 Sequence 3
16	2203	63.8	3475	6	144733	144733 Sequence 3
17	2203	63.8	3475	9	HSU02687	U02687 Human growt
18	2191.4	63.5	3422	9	HSFLT3RTK	Z26652 H. sapiens F
19	2081.8	60.3	3120	6	AR262794	AR262794 Sequence
20	2081.8	60.3	3120	6	144518	144518 Sequence 22
21	2069.8	59.9	2982	6	AX695523	AX695523 Sequence
22	1206	34.9	1894	6	AR149571	AR149571 Sequence
23	1119.6	32.4	2247	6	BD022236	BD022236 Multi-fun
24	474.6	13.7	152564	10	AC127549	AC127549 Mus muscu
25	470.6	13.6	267290	2	AC134441	AC134441 Mus muscu
26	427.4	12.4	429	10	MUS3RTK	L36163 Mus musculu
27	363.2	10.5	480	9	HUM3RTK	L36162 Homo sapien
28	357.4	10.4	3992	6	AX394288	AX394288 Sequence
29	357.4	10.4	3992	6	AX587822	AX587822 Sequence
30	357.4	10.4	3992	9	HSCFMS	X03663 Human mRNA
31	355.8	10.3	3904	9	BC047521	BC047521 Homo sapi
32	349.8	10.1	3545	5	AF153446	AF153446 Danio rer
33	347	10.0	4588	5	OMY417832	AJ417832 Oncorhyn
34	344.2	10.0	389	10	AY094358	AY094358 Rattus no
35	341.8	9.9	4577	14	FCSSMONC	X01643 Feline sarc
36	340.2	9.9	3828	4	CATFMS	J03149 Cat (F dome
37	336.8	9.8	4222	4	S76596	S76596 C-kit-recep
38	335.6	9.7	3705	10	RRCKITRTK	X62491 R.rattus mr
39	334.4	9.7	266001	2	AC106082	AC106082 Rattus no
40	333.6	9.7	3691	10	BC043054	BC043054 Mus muscu
41	333.6	9.7	3699	10	BC036343	BC036343 Mus muscu
42	332.6	9.6	2176	4	AF263826	AF263826 Bos tauru
43	332.6	9.6	2176	4	AF263827	AF263827 Bos tauru
44	332.6	9.6	3069	4	BOVCKR	D16680 Bovine mRNA
45	331.4	9.6	2919	6	A82459	A82459 Sequence 47

ALIGNMENTS

RESULT 1

AR005211

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

AR005211

Sequence 1 from patent US 5747651.

AR005211

AR005211.1

GI:3966090

Unknown.

Unknown.

1 (bases 1 to 3453)

Lemischka, I.R.

Antibodies against tyrosine kinase receptor flk-1

Patent: US 5747651-A 1 05-MAY-1998;

US 5747651-A 1 05-MAY-1998;

Location/Qualifiers

3453 bp

DNA

linear

PAT 04-DEC-1998

Pred. No. is the number of results predicted by chance to have a

[illegible]

2041	Db		CACCATGACAACATCGTGAATCTGCTGGGGCATGCACACTGTCAGGCGCAGTGTACTTG	2100
2101	Qy		ATTTTTGAATATTGTTGCTATGCTGACCTCTCAACTACTCTAAGAAGTAAAGAGAGAAG	2160
2101	Db			2160
2161	Qy		TTTTCACAGGACATGACAGAGATTTTTTAAGGAACAATAATTTCAGTTCCTTACCTACTTTC	2220
2161	Db		TTTTCACAGGACATGACAGAGATTTTTTAAGGAACAATAATTTCAGTTCCTTACCTACTTTC	2220
2221	Qy		CAGGCAATTCAAATTCACAGATGCCTGGTTACAGAGAAGTTTCAGTTACACCCGCCCTTG	2280
2221	Db		CAGGCAATTCAAATTCACAGATGCCTGGTTACAGAGAAGTTTCAGTTACACCCGCCCTTG	2280
2281	Qy		GATCAGCTCTCAGGTTCAATGGGAATTCAAATTCCTGGAAGTGCAGATTGAAATGAA	2340
2281	Db		GATCAGCTCTCAGGTTCAATGGGAATTCAAATTCCTGGAAGTGCAGATTGAAATGAA	2340
2341	Qy		AACCAAGAAGGCTGGCAGAAAGAAGAGGAAGAATTTGAAAGTGTCTGACGTTTCAAGAC	2400
2341	Db		AACCAAGAAGGCTGGCAGAAAGAAGAGGAAGAATTTGAAAGTGTCTGACGTTTCAAGAC	2400
2401	Qy		CTCCTTTGCTTTGCGTACCAAGTGGCCAAAGGCATGGAAATTCCTCGAGTTCAGTCTGCT	2460
2401	Db		CTCCTTTGCTTTGCGTACCAAGTGGCCAAAGGCATGGAAATTCCTCGAGTTCAGTCTGCT	2460
2461	Qy		GTTCCACAGAGACTGGCAGCCAGGAATGTGTTGGTCAACCACCGGAAGTGTGAGATC	2520
2461	Db		GTTCCACAGAGACTGGCAGCCAGGAATGTGTTGGTCAACCACCGGAAGTGTGAGATC	2520
2521	Qy		TGTGACTTTGGAATGCGCCGAGACATCCTCAGCGACTTCAGCTACGTCGTCAGGGGCAAC	2580
2521	Db		TGTGACTTTGGAATGCGCCGAGACATCCTCAGCGACTTCAGCTACGTCGTCAGGGGCAAC	2580
2581	Qy		GCACGGCTGCCGGTGAAGTGATGGCACCCGAGAGCTTATTGGAAGGATCTACACAATC	2640
2581	Db		GCACGGCTGCCGGTGAAGTGATGGCACCCGAGAGCTTATTGGAAGGATCTACACAATC	2640
2641	Qy		AAGAGTGAACGTCTGGTCTCTACGGCATCCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC	2700
2641	Db		AAGAGTGAACGTCTGGTCTCTACGGCATCCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC	2700
2701	Qy		CCTTACCCTGGCATTCCTGTGACGCTAACTTTATATAACTGATTACAGATGATTTAA	2760
2701	Db		CCTTACCCTGGCATTCCTGTGACGCTAACTTTATATAACTGATTACAGATGATTTAA	2760
2761	Qy		ATGGAGACCATTCATATGCCACAGAGGGATATACTTTGTATATGCATCCTCTGGGCT	2820
2761	Db		ATGGAGACCATTCATATGCCACAGAGGGATATACTTTGTATATGCATCCTCTGGGCT	2820
2821	Qy		TTTGTACTCAAGAGCGGCCATCTCTCCCAACTGACTTCAATTTTAGGATGTACGCTG	2880
2821	Db		TTTGTACTCAAGAGCGGCCATCTCTCCCAACTGACTTCAATTTTAGGATGTACGCTG	2880
2881	Qy		GCAGAGGACAGAGAAGCATGTATCAGAACATCCATCCATTCACAAACAGGGGGCCCT	2940
2881	Db		GCAGAGGACAGAGAAGCATGTATCAGAACATCCATCCATTCACAAACAGGGGGCCCT	2940
2941	Qy		CAGCAGAGAGCGGCTCAGAGCCAGTGCACACGCGCCAGGTGAAGATTTCACAGAA	3000
2941	Db		CAGCAGAGAGCGGCTCAGAGCCAGTGCACACGCGCCAGGTGAAGATTTCACAGAA	3000
3001	Qy		AGAGTTACGAGAGGCGCTTGGACCCGCCACCTAGCAGGCTGTAGACCGCAGAGCCA	3060
3001	Db		AGAGTTACGAGAGGCGCTTGGACCCGCCACCTAGCAGGCTGTAGACCGCAGAGCCA	3060
3061	Qy		AGATTAGCCTCGCTCTGAGGAAGCGCCCTACAGCGCTTGTCTGCTGGATCTTCTCT	3120
3061	Db		AGATTAGCCTCGCTCTGAGGAAGCGCCCTACAGCGCTTGTCTGCTGGATCTTCTCT	3120
3121	Qy		AGATGCTGTCTGCCATTCTCCAAAGTGACTTCTATAAAATCAAAACCTCTCCTCGCAG	3180

[illegible]

Db 361 AGCTCCCTGGGCTGCCAGCCGCACTTTGATTTTACAAAACAGAGGAATCGTTTCCATGGCC 420  
 Qy 421 ATCTTGAACTGTACAGAGACCCAGGACAGGAGAAATACCTACTTCCATATTTACAGGCAAGC 480  
 Db 421 ATCTTGAACTGTACAGAGACCCAGGACAGGAGAAATACCTACTTCCATATTTACAGGCAAGC 480  
 Qy 481 GCCAACTACACAGTACTGTTTACAGTGAAATGTAAGAGATACACAGCTGTATGTCTAAGG 540  
 Db 481 GCCAACTACACAGTACTGTTTACAGTGAAATGTAAGAGATACACAGCTGTATGTCTAAGG 540  
 Qy 541 AGACCTTACTTTAGGAAGATGGAAGAACCCAGGATGCACTGCTCTGCACTCCGAGGGTGT 600  
 Db 541 AGACCTTACTTTAGGAAGATGGAAGAACCCAGGATGCACTGCTCTGCACTCCGAGGGTGT 600  
 Qy 601 CCGGAGCCCACTGTGAGTGGGTCTCTGAGCTCCACAGGAAAGAGCTGTAAAGAGAA 660  
 Db 601 CCGGAGCCCACTGTGAGTGGGTCTCTGAGCTCCACAGGAAAGAGCTGTAAAGAGAA 660  
 Qy 661 GGCCTCTGCTGTGTACAGAAAGAGGAGAAAGGTACTTCACTAGTGTGTTGCGAAACAGACATC 720  
 Db 661 GGCCTCTGCTGTGTACAGAAAGAGGAGAAAGGTACTTCACTAGTGTGTTGCGAAACAGACATC 720  
 Qy 721 AGATGCTGTGTAGAAATGCACTGGGCGCGGAATGCAACCAAGCTGTTTCCACATAGATCTA 780  
 Db 721 AGATGCTGTGTAGAAATGCACTGGGCGCGGAATGCAACCAAGCTGTTTCCACATAGATCTA 780  
 Qy 781 AACCAGGCTCTCAGAGCACACTGCCCCAGTATTTCTGAAAGTGGGGAGACCCCTGTGG 840  
 Db 781 AACCAGGCTCTCAGAGCACACTGCCCCAGTATTTCTGAAAGTGGGGAGACCCCTGTGG 840  
 Qy 841 ATCAGGTGTAAGGCCATCCATGTGAACCATGGATTCGGGCTCACCTGGGAGCTGGAAGAC 900  
 Db 841 ATCAGGTGTAAGGCCATCCATGTGAACCATGGATTCGGGCTCACCTGGGAGCTGGAAGAC 900  
 Qy 901 AAGGCCCTGGAGGAGGAGCTACTTTGAGATGAGTACCTTACTTCCACAAACAGGACCATG 960  
 Db 901 AAGGCCCTGGAGGAGGAGCTACTTTGAGATGAGTACCTTACTTCCACAAACAGGACCATG 960  
 Qy 961 ATTCGGATCTCTTGGCTTTGCTCTTCCGTGGGAGGAAACGACACCGGATATATACCC 1020  
 Db 961 ATTCGGATCTCTTGGCTTTGCTCTTCCGTGGGAGGAAACGACACCGGATATATACCC 1020  
 Qy 1021 TGCTCTTCTCAAAGCACCCAGCCAGTCAAGGCTTGGTGACCATCTCTAGAAAAGGGTTT 1080  
 Db 1021 TGCTCTTCTCAAAGCACCCAGCCAGTCAAGGCTTGGTGACCATCTCTAGAAAAGGGTTT 1080  
 Qy 1081 ATAAACGCTTACAGCTCGCAAGAGATGAAATTTGACCCGTACGAAAGTTCTGCTTC 1140  
 Db 1081 ATAAACGCTTACAGCTCGCAAGAGATGAAATTTGACCCGTACGAAAGTTCTGCTTC 1140  
 Qy 1141 TCAGTCAAGCTTAAAGCTTACCAAGATCCGATGCACTGATCTTCTCAAGCCTCA 1200  
 Db 1141 TCAGTCAAGCTTAAAGCTTACCAAGATCCGATGCACTGATCTTCTCAAGCCTCA 1200  
 Qy 1201 TTTCTCTGTGAACAGAGAGGCTGGAGGATGGGTACAGATATCTTAAATTTTGGCATCAT 1260  
 Db 1201 TTTCTCTGTGAACAGAGAGGCTGGAGGATGGGTACAGATATCTTAAATTTTGGCATCAT 1260  
 Qy 1261 AAGAAACAGCCAGGAGAGTACATATTTCTATGAGAAATGATGAGCCCGAGTTTCCACAAA 1320  
 Db 1261 AAGAAACAGCCAGGAGAGTACATATTTCTATGAGAAATGATGAGCCCGAGTTTCCACAAA 1320  
 Qy 1321 ATGTTTCAAGCTGAAATATGAAGAGAAACCTCAAGTGTAGCAAAATGCTTACGCCAGCAG 1380  
 Db 1321 ATGTTTCAAGCTGAAATATGAAGAGAAACCTCAAGTGTAGCAAAATGCTTACGCCAGCAG 1380  
 Qy 1381 GCGTCTCTCTCTGATGGCTTACCCGCTACCTCTGGAAGGAGTGTGGAATGGAAGTGTGGAC 1440  
 Db 1381 GCGTCTCTCTCTGATGGCTTACCCGCTACCTCTGGAAGGAGTGTGGAATGGAAGTGTGGAC 1440  
 Qy 1441 AAATCTCCAAATGTCAGGAGGAAATCCAGAGGAGTGTGGAATGGAAGGCTTAACAGA 1500  
 Db 1441 AAATCTCCAAATGTCAGGAGGAAATCCAGAGGAGTGTGGAATGGAAGGCTTAACAGA 1500

Qy 1501 AAAGTGTGTTGGCCAGTGGGTGTGAGCACTCTAAATATGATGAGGCGGGAAGGG 1560  
 Db 1501 AAAGTGTGTTGGCCAGTGGGTGTGAGCACTCTAAATATGATGAGGCGGGAAGGG 1560  
 Qy 1561 CTTCTGTCAAAATGCTGTGCTCAATTTCTATGGGCAAGCTTGGGAAACCATCTTTTAA 1620  
 Db 1561 CTTCTGTCAAAATGCTGTGCTCAATTTCTATGGGCAAGCTTGGGAAACCATCTTTTAA 1620  
 Qy 1621 AACTCACAGGCGCCCTTCCCTTTTCAATCAACAGAAACAACTCTCTTCTATGCAACCATTTGG 1680  
 Db 1621 AACTCACAGGCGCCCTTCCCTTTTCAATCAACAGAAACAACTCTCTTCTATGCAACCATTTGG 1680  
 Qy 1681 CTCTGTCTCCCTTCAATTTGTTGTTCTCATTTGTTGATCTGCTTCTTATGGACCATTTGG 1740  
 Db 1681 CTCTGTCTCCCTTCAATTTGTTGTTCTCATTTGTTGATCTGCTTCTTATGGACCATTTGG 1740  
 Qy 1741 TTTAGGTACGAGAGTCAAGTGCAGATGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 1800  
 Db 1741 TTTAGGTACGAGAGTCAAGTGCAGATGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 1800  
 Qy 1801 TTCTACGTTGCACTTCAAGGAGTATGAAATATGACCTTAAATGGGAGTTCCTCCAGAGAGAAC 1860  
 Db 1801 TTCTACGTTGCACTTCAAGGAGTATGAAATATGACCTTAAATGGGAGTTCCTCCAGAGAGAAC 1860  
 Qy 1861 TTAGATTTGGGAGGTCTCGGCTCTGGGCTCTGGGCTCTGGGCTCTGGGCTCTGGGCTCTGGGCTCT 1920  
 Db 1861 TTAGATTTGGGAGGTCTCGGCTCTGGGCTCTGGGCTCTGGGCTCTGGGCTCTGGGCTCTGGGCTCT 1920  
 Qy 1921 TATGGCATTTAGTAAACCGGAGTCTCAATTTCAAGTGGCGGTCTGGGAGGTGATGAAACGCAACG 1980  
 Db 1921 TATGGCATTTAGTAAACCGGAGTCTCAATTTCAAGTGGCGGTCTGGGAGGTGATGAAACGCAACG 1980  
 Qy 1981 GCTGACAGCTGTGAAAGAGAGCTCTCATGTGGAGCTCAAAATGATGACCCACCTGGGA 2040  
 Db 1981 GCTGACAGCTGTGAAAGAGAGCTCTCATGTGGAGCTCAAAATGATGACCCACCTGGGA 2040  
 Qy 2041 CACCATGACAAATCGTGAATCTGCTGGGCGCATGCACTGTCAAGGCGAGTGTACTTG 2100  
 Db 2041 CACCATGACAAATCGTGAATCTGCTGGGCGCATGCACTGTCAAGGCGAGTGTACTTG 2100  
 Qy 2101 ATTTTGAATATTTGTTGATGTTTAAAGGACATAATTTTAAAGGACATAATTTTAAAGGACATA 2160  
 Db 2101 ATTTTGAATATTTGTTGATGTTTAAAGGACATAATTTTAAAGGACATAATTTTAAAGGACATA 2160  
 Qy 2161 TTTTCAAGGAGCTGCGAG 2220  
 Db 2161 TTTTCAAGGAGCTGCGAG 2220  
 Qy 2221 CAGGCACATTTCAAAATTTCCAGCATGCTGCTTCCAGGAGAGTTCAGTTTACACCCCGCTTTC 2280  
 Db 2221 CAGGCACATTTCAAAATTTCCAGCATGCTGCTTCCAGGAGAGTTCAGTTTACACCCCGCTTTC 2280  
 Qy 2281 GATCAGCTCTCAGGTTTCAATGGGAAATTTCAATTTCTTGAAGATGAGATTTGAATATGAA 2340  
 Db 2281 GATCAGCTCTCAGGTTTCAATGGGAAATTTCAATTTCTTGAAGATGAGATTTGAATATGAA 2340  
 Qy 2341 AACCAAGAGAGCTGCGAG 2400  
 Db 2341 AACCAAGAGAGCTGCGAG 2400  
 Qy 2401 CTCCTTTGCTTTGCTTACCAAGTGGCCCAAGGAGTGGAAATTTCTTGAAGTTCAGTCTGCTGT 2460  
 Db 2401 CTCCTTTGCTTTGCTTACCAAGTGGCCCAAGGAGTGGAAATTTCTTGAAGTTCAGTCTGCTGT 2460  
 Qy 2461 GTCCACAGAGACCTGGCAGCAG 2520  
 Db 2461 GTCCACAGAGACCTGGCAGCAG 2520  
 Qy 2521 TGTGATTTTGAAGTGGCCCGGAG 2580  
 Db 2521 TGTGATTTTGAAGTGGCCCGGAG 2580





Db 901 AAAGCCCTGGAGGAGGCGAGCTACTTGGATGAGTACCTACTCCACAAAGAGCAACATG 960  
Qy 961 ATTTCGATTCCTCTGSCCTTTGTCTTCCTCGTGGGAGGAAAGACACCGGATATACACC 1020  
Db 961 ATTTCGATTCCTCTGSCCTTTGTCTTCCTCGTGGGAGGAAAGACACCGGATATACACC 1020  
Qy 1021 TGCTCTTCTCAAAGCACCCAGCCAGTGGGTGGTACCATCTCTAGAAAAGGGTTT 1080  
Db 1021 TGCTCTTCTCAAAGCACCCAGCCAGTGGGTGGTACCATCTCTAGAAAAGGGTTT 1080  
Qy 1081 ATAAACGCTACAGCTCGCAAGAGATGAATTTGACCCGTACGAAAGTTCTGCTTC 1140  
Db 1081 ATAAACGCTACAGCTCGCAAGAGATGAATTTGACCCGTACGAAAGTTCTGCTTC 1140  
Qy 1141 TCAGTCAGCTTTAAAGCGTACCCAGATCCGATCCGATGCAATCTCTCAAGCCTCA 1200  
Db 1141 TCAGTCAGCTTTAAAGCGTACCCAGATCCGATCCGATGCAATCTCTCAAGCCTCA 1200  
Qy 1201 TTTCTCTGTGAACAGAGAGGCTGGAGATGGGTACAGCATATCTAAATTTTTCGATCAT 1260  
Db 1201 TTTCTCTGTGAACAGAGAGGCTGGAGATGGGTACAGCATATCTAAATTTTTCGATCAT 1260  
Qy 1261 AAGAACAGCCAGGAGGTACATATCTATGACAGAAATGATGACGCCAGTTTCAACAAA 1320  
Db 1261 AAGAACAGCCAGGAGGTACATATCTATGACAGAAATGATGACGCCAGTTTCAACAAA 1320  
Qy 1321 ATGTTTACGCTGAATATGAAGAGAACTCAAGTGTCTAGCAATGCTCAGCCAGCCAG 1380  
Db 1321 ATGTTTACGCTGAATATGAAGAGAACTCAAGTGTCTAGCAATGCTCAGCCAGCCAG 1380  
Qy 1381 GCGTCTCTGTTCTCTGATGGCTACCCGCTACCCCTCTGACCTGGAAGAGTGTTCGGAC 1440  
Db 1381 GCGTCTCTGTTCTCTGATGGCTACCCGCTACCCCTCTGACCTGGAAGAGTGTTCGGAC 1440  
Qy 1441 AAATCTCCCAATGACAGAGGAAATCCAGAGGAGTTTGGAAATAAAGAGCTTAACAGA 1500  
Db 1441 AAATCTCCCAATGACAGAGGAAATCCAGAGGAGTTTGGAAATAAAGAGCTTAACAGA 1500  
Qy 1501 AAAGTGTTCGCCATGCTGGTTCAGAGTACTCTAAATATGATGAGGCGCGGAAAGGG 1560  
Db 1501 AAAGTGTTCGCCATGCTGGTTCAGAGTACTCTAAATATGATGAGTGAAGCGCGGAAAGGG 1560  
Qy 1561 CTTCTGTCTCAAATGCTGTGCTGACAAATCTATGCGCAGCTCTGGAACCATCTTTTAA 1620  
Db 1561 CTTCTGTCTCAAATGCTGTGCTGACAAATCTATGCGCAGCTCTGGAACCATCTTTTAA 1620  
Qy 1621 AACTCAGCGGCCCTTCCCTTTTCAATCCAGACAAACATCTCCTTCTATGCGACATTTGGG 1680  
Db 1621 AACTCAGCGGCCCTTCCCTTTTCAATCCAGACAAACATCTCCTTCTATGCGACATTTGGG 1680  
Qy 1681 CTCTGTCTCCCTTCATTTGTTCTCATTTGTTGATCTGCGCAAAATACAAAAGCAA 1740  
Db 1681 CTCTGTCTCCCTTCATTTGTTCTCATTTGTTGATCTGCGCAAAATACAAAAGCAA 1740  
Qy 1741 TTTAGGTACAGAGTACAGTGCAGATCATCCAGTGCATCGCCCCCTGGATACAGGATAC 1800  
Db 1741 TTTAGGTACAGAGTACAGTGCAGATCATCCAGTGCATCGCCCCCTGGATACAGGATAC 1800  
Qy 1801 TTCTACGTTGACTTCAGGACTATGAATATGACCTTAAGTGGGATTCCTCCAGAGAGAAC 1860  
Db 1801 TTCTACGTTGACTTCAGGACTATGAATATGACCTTAAGTGGGATTCCTCCAGAGAGAAC 1860  
Qy 1861 TTAGAGTTTGGAGGTCCTGGGGTCTGGGGTCTGGGGTCTGGGGTCTGGGGTCTGGGGT 1920  
Db 1861 TTAGAGTTTGGAGGTCCTGGGGTCTGGGGTCTGGGGTCTGGGGTCTGGGGTCTGGGGT 1920  
Qy 1921 TATGGCATTTAGTAAACCGGAGTCTCAATTCAGTGGCGGTGAAGTCTTAAAGAGAAA 1980  
Db 1921 TATGGCATTTAGTAAACCGGAGTCTCAATTCAGTGGCGGTGAAGTCTTAAAGAGAAA 1980  
Qy 1981 GCTGACAGCTGTGAAAAGAGAGCTCTATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2040  
Db 1981 GCTGACAGCTGTGAAAAGAGAGCTCTATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2040

Db 1981 GCTGACAGCTGTGAAAAGAGAGCTCTATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2040  
Qy 2041 CACCATGACAAACATCGTGAATCTGTGGGGGATGCACACTGTGAGGGCCAGTGTACTTG 2100  
Db 2041 CACCATGACAAACATCGTGAATCTGTGGGGGATGCACACTGTGAGGGCCAGTGTACTTG 2100  
Qy 2101 ATTTTGTGAATATTTGTTGCTATGCTGACCTCTCACTACCTAAGAAAGTAAAGAGAGAAG 2160  
Db 2101 ATTTTGTGAATATTTGTTGCTATGCTGACCTCTCACTACCTAAGAAAGTAAAGAGAGAAG 2160  
Qy 2161 TTTTTCAGGACATGACAGAGATTTTAAAGGAACATAATTTTCAAGTCTTACCTACTTTC 2220  
Db 2161 TTTTTCAGGACATGACAGAGATTTTAAAGGAACATAATTTTCAAGTCTTACCTACTTTC 2220  
Qy 2221 CAGGCACATTTCAAATTTCCAGCATGCTGTTTCAAGAGAGTTTCAAGTACACCGCCCTTG 2280  
Db 2221 CAGGCACATTTCAAATTTCCAGCATGCTGTTTCAAGAGAGTTTCAAGTACACCGCCCTTG 2280  
Qy 2281 GATCAGCTCTCAGGGTTCAAATGGGAAATTTCAATTTCAATCTTGAAGATGAGATTGAATGAA 2340  
Db 2281 GATCAGCTCTCAGGGTTCAAATGGGAAATTTCAATTTCAATCTTGAAGATGAGATTGAATGAA 2340  
Qy 2341 AACCAAGAGAGGCTGGCAG 2400  
Db 2341 AACCAAGAGAGGCTGGCAG 2400  
Qy 2401 CTCTCTTGTCTTGTGCTACCAAGTGGCCAAAGGCAATGTTTGTGTGTGTGTGTGTGTGTGT 2460  
Db 2401 CTCTCTTGTCTTGTGCTACCAAGTGGCCAAAGGCAATGTTTGTGTGTGTGTGTGTGTGTGT 2460  
Qy 2461 GTCCACAGAGACCTGGGACAGGAGATGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2520  
Db 2461 GTCCACAGAGACCTGGGACAGGAGATGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2520  
Qy 2521 TGTGACCTTGTGACCTGGGCGGAGACATCTGTGAGGACCTCCAGTACCTGTCAGGGGCAAC 2580  
Db 2521 TGTGACCTTGTGACCTGGGCGGAGACATCTGTGAGGACCTCCAGTACCTGTCAGGGGCAAC 2580  
Qy 2581 GCACGGCTGCGGCTGAAGTGGATGGCAGCCGAGAGCTTATTTGAAGGATCTACACAATC 2640  
Db 2581 GCACGGCTGCGGCTGAAGTGGATGGCAGCCGAGAGCTTATTTGAAGGATCTACACAATC 2640  
Qy 2641 AAGTGTGACCTGCTGCTCCTACGGCATCTTCTCTGAGAGATATTTTCACTGGGTGTGAAC 2700  
Db 2641 AAGTGTGACCTGCTGCTCCTACGGCATCTTCTCTGAGAGATATTTTCACTGGGTGTGAAC 2700  
Qy 2701 CCTTACCTGCGCATTCCTGTGAGCGCTAACTTCTATATAAATGATTCAGAGTGAATTTAAA 2760  
Db 2701 CCTTACCTGCGCATTCCTGTGAGCGCTAACTTCTATATAAATGATTCAGAGTGAATTTAAA 2760  
Qy 2761 ATGGAGCAGCCATTTCTATGCCACAGAGGATATACCTTTGTAATGCAATCTCTGCTGGGCT 2820  
Db 2761 ATGGAGCAGCCATTTCTATGCCACAGAGGATATACCTTTGTAATGCAATCTCTGCTGGGCT 2820  
Qy 2821 TTTGACCTCAAGGAAGCGGCCATCTTCCCAACCTGACTTCAATTTTGAAGTGTGAGCTG 2880  
Db 2821 TTTGACCTCAAGGAAGCGGCCATCTTCCCAACCTGACTTCAATTTTGAAGTGTGAGCTG 2880  
Qy 2881 GCAGAGCAGAGAGAGAGATGATACAGACATCCATCTACAAAACAGGGCGGCCCT 2940  
Db 2881 GCAGAGCAGAGAGAGAGATGATACAGACATCCATCTACAAAACAGGGCGGCCCT 2940  
Qy 2941 CAGCAGAGGCGGGCTCAGAGCCAGTCCAGAGCGCCAGCGGAGAGATTTCAAGAGAA 3000  
Db 2941 CAGCAGAGGCGGGCTCAGAGCCAGTCCAGAGCGCCAGCGGAGAGATTTCAAGAGAA 3000  
Qy 3001 AGAAGTTACGAGAGGCGCTTGGACCCCGCCACCTTAGCAGGCTGTAGACCGCAGAGCCA 3060  
Db 3001 AGAAGTTACGAGAGGCGCTTGGACCCCGCCACCTTAGCAGGCTGTAGACCGCAGAGCCA 3060  
Qy 3061 AGATTAGCCTCGCCTCTGAGAGAGCGCCCTACAGCGCGTGTCTCGCTGAGACTTTTCTCT 3120  
Db 3061 AGATTAGCCTCGCCTCTGAGAGAGCGCCCTACAGCGCGTGTCTCGCTGAGACTTTTCTCT 3120





Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE	1
AUTHORS	Morris, D.W. and Engelhard, E.K.
TITLE	Novel compositions and methods for cancer
JOURNAL	Patent: WO 03008583-A 1146 30-JAN-2003;
	Sagres Discovery (US)
FEATURES	Location/Qualifiers
source	1..3426
	/organism="Mus musculus"
	/mol_type="genomic DNA"
	/db_xref="taxon:10090"
BASE COUNT	920 a 822 c 875 g 809 t
ORIGIN	
Query Match	99.2%; Score 3426; DB 6; Length 3426;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 3426; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 GCGGCTGGCTACCGCGCGTCCGGAGCCATCGGGCGTTGGCGGAGCGCAGCGACCGG 60
Db	1 GCGGCTGGCTACCGCGCGTCCGGAGCCATCGGGCGTTGGCGGAGCGCAGCGACCGG 60
Qy	61 CGGCTGCTGCTGTGTTTGTTCAGTAATGATTTCTTGAGACCGTTACAAACCAAGAC 120
Db	61 CGGCTGCTGCTGTGTTTGTTCAGTAATGATTTCTTGAGACCGTTACAAACCAAGAC 120
Qy	121 CTGCTGTGATCAAGTGTGTTTTAATCAGTCATGAGAAACAATGCTCATCAGCGGGAAAG 180
Db	121 CTGCTGTGATCAAGTGTGTTTTAATCAGTCATGAGAAACAATGCTCATCAGCGGGAAAG 180
Qy	181 CCATCATCTGTAACCGAATGGTTCGAGGATCCCCAGAGACCTTCCAGTGTAACCCGAGCGC 240
Db	181 CCATCATCTGTAACCGAATGGTTCGAGGATCCCCAGAGACCTTCCAGTGTAACCCGAGCGC 240
Qy	241 CAGAGTGAAGGGAACGGTATATGAAGCGGCCAACCGTGAAGTGGCCGAGTCTGGGTCCATC 300
Db	241 CAGAGTGAAGGGAACGGTATATGAAGCGGCCAACCGTGAAGTGGCCGAGTCTGGGTCCATC 300
Qy	301 ACCCTGCAAGTGCAGCTCGCCACCCAGGGGACCTTTCTGCCTCTGGGTCTTTAAGCAC 360
Db	301 ACCCTGCAAGTGCAGCTCGCCACCCAGGGGACCTTTCTGCCTCTGGGTCTTTAAGCAC 360
Qy	361 AGCTCCCTGGGCTGCAGCGCCGACCTTTGATTTTACAAAACAGAGGAATCGTTTCCATGCC 420
Db	361 AGCTCCCTGGGCTGCAGCGCCGACCTTTGATTTTACAAAACAGAGGAATCGTTTCCATGCC 420
Qy	421 ATCTTTGAACGTTGACAGAGACCCAGGAGGAATACCTACTCCATATTTAGAGCGGACGC 480
Db	421 ATCTTTGAACGTTGACAGAGACCCAGGAGGAATACCTACTCCATATTTAGAGCGGACGC 480
Qy	481 GCCAACTACAGTACTGTTTCAGTGAATGTAAGAGATACACAGCTGTATGTGCTAAGG 540
Db	481 GCCAACTACAGTACTGTTTCAAGTGAATGTAAGAGATACACAGCTGTATGTGCTAAGG 540
Qy	541 AGACCTTACTTTAGGAAGATGGAACCAAGATGCACTGTCTGTCAATCTCCGAGGGTGT 600
Db	541 AGACCTTACTTTAGGAAGATGGAACCAAGATGCACTGTCTGTCAATCTCCGAGGGTGT 600
Qy	601 CCGGAGCCCACTGTGTGAGTGGGTGCTCTGACGCTCCACAGGGAAGCTGTAAAGAA 660
Db	601 CCGGAGCCCACTGTGTGAGTGGGTGCTCTGACGCTCCACAGGGAAGCTGTAAAGAA 660
Qy	661 GGGCCCTGCTGTTGTCAAGAGGGGAAAAGGTACTTTCATGAGTTGTTTCGAAACAGACATC 720
Db	661 GGGCCCTGCTGTTGTCAAGAGGGGAAAAGGTACTTTCATGAGTTGTTTCGAAACAGACATC 720
Qy	721 AGATGCTGTGCTAGAAATGCACCTGGGCCCGGAATGCACCAAGCTGTTCACCATAGATCTA 780
Db	721 AGATGCTGTGCTAGAAATGCACCTGGGCCCGGAATGCACCAAGCTGTTCACCATAGATCTA 780
Qy	781 AACCAAGGCTCCTCAGAGCACTGCCCCAGTTATTCCTGAAGTGGGGAAACCTTGTGG 840

QY 1921 TATGCAATTAGTAAACCGGAGTCTCAATTCAGGTGGCGTGAAGATGCTAAAGAGAAA 1980  
DB 1921 TATGCAATTAGTAAACCGGAGTCTCAATTCAGGTGGCGTGAAGATGCTAAAGAGAAA 1980  
QY 1981 GCTGACAGCTGTGAAAGAAAGCTCTCATGTCGGAAGCTCAAAATGATGACCCACCTGGGA 2040  
DB 1981 GCTGACAGCTGTGAAAGAAAGCTCTCATGTCGGAAGCTCAAAATGATGACCCACCTGGGA 2040  
QY 2041 CACCATGACAAATCTGTGATCTGCTGGGGGATGACACATGTCAGGGCCAGTGTACTTG 2100  
DB 2041 CACCATGACAAATCTGTGATCTGCTGGGGGATGACACATGTCAGGGCCAGTGTACTTG 2100  
QY 2101 ATTTTGAATATTTGTTCTATGTCGACCTCTCAATCTACCTAAGAACTAAAGAGAGAG 2160  
DB 2101 ATTTTGAATATTTGTTCTATGTCGACCTCTCAATCTACCTAAGAACTAAAGAGAGAG 2160  
QY 2161 TTTCAGACATGGAAGAGATTTTAAAGAACTAAATTTCAAGTCTTTACCTACTTTTC 2220  
DB 2161 TTTCAGACATGGAAGAGATTTTAAAGAACTAAATTTCAAGTCTTTACCTACTTTTC 2220  
QY 2221 CAGGCACATTCAAATTCAGCATGCTGTTCAAGAGAGTTCAGTTACACCGCCCTTG 2280  
DB 2221 CAGGCACATTCAAATTCAGCATGCTGTTCAAGAGAGTTCAGTTACACCGCCCTTG 2280  
QY 2281 GATCAGCTCTCAGGGTTCAATGGGAATTCATTTCAATCTGAAGATGAGATTTGAATGAA 2340  
DB 2281 GATCAGCTCTCAGGGTTCAATGGGAATTCATTTCAATCTGAAGATGAGATTTGAATGAA 2340  
QY 2341 AACGAGAGAGCTGCGACAGAGAGAGAGAGATTTGAAGCTGCTGACGTTTGAAGAC 2400  
DB 2341 AACGAGAGAGCTGCGACAGAGAGAGAGAGATTTGAAGCTGCTGACGTTTGAAGAC 2400  
QY 2401 CTCCTTTGCTTTGCTGACCAAGTGGCCAAAGGATTCCTGGAGTTCAAGTCGTGT 2460  
DB 2401 CTCCTTTGCTTTGCTGACCAAGTGGCCAAAGGATTCCTGGAGTTCAAGTCGTGT 2460  
QY 2461 GTCCACAGAGACTGCGACAGAGAGATTTGTTGTTCAACCCAGGAGAGTGTGAGATC 2520  
DB 2461 GTCCACAGAGACTGCGACAGAGAGATTTGTTGTTCAACCCAGGAGAGTGTGAGATC 2520  
QY 2521 TGTGACTTTGGACTGCGCCGAGACATCTGAGCGACTCCAGCTAGCTGCTGAGGGGCAAC 2580  
DB 2521 TGTGACTTTGGACTGCGCCGAGACATCTGAGCGACTCCAGCTAGCTGCTGAGGGGCAAC 2580  
QY 2581 GCAAGCTGCGCGTGAAGTGGATGCAACCGAGAGCTTTATTTGAAGGGATCTACAAATC 2640  
DB 2581 GCAAGCTGCGCGTGAAGTGGATGCAACCGAGAGCTTTATTTGAAGGGATCTACAAATC 2640  
QY 2641 AAGAGTGAAGCTGCTGCTGACGGCATCTTCTGCGGAGATATTTTCACTGGGTGTGAAC 2700  
DB 2641 AAGAGTGAAGCTGCTGCTGACGGCATCTTCTGCGGAGATATTTTCACTGGGTGTGAAC 2700  
QY 2701 CTTTACCTGCGATTCGTGCGACCTAACTTCTATAAAGCTGATTCAGAGTGGATTTAAA 2760  
DB 2701 CTTTACCTGCGATTCGTGCGACCTAACTTCTATAAAGCTGATTCAGAGTGGATTTAAA 2760  
QY 2761 ATGGAGAGGCAATTTATGCGCAGAGAGGATATCTTTGTAATGCAATCTGCTGGGCT 2820  
DB 2761 ATGGAGAGGCAATTTATGCGCAGAGAGGATATCTTTGTAATGCAATCTGCTGGGCT 2820  
QY 2821 TTTGACTCAAGGAGCGGCTCCTTTCCCAACCTGACTTCATTTTATAGGATGTGAGCTG 2880  
DB 2821 TTTGACTCAAGGAGCGGCTCCTTTCCCAACCTGACTTCATTTTATAGGATGTGAGCTG 2880  
QY 2881 GCAGGAGAGAGAGCAATGTATCAGAAATCTCCATCCATCCATCCAAACAGCGCGCCCT 2940  
DB 2881 GCAGGAGAGAGAGCAATGTATCAGAAATCTCCATCCATCCATCCAAACAGCGCGCCCT 2940  
QY 2941 CAGCAGAGAGCGGGCTCAGAGCCAGTCCGACAGCGCCAGGTGAAGATTCACAGAGAA 3000  
DB 2941 CAGCAGAGAGCGGGCTCAGAGCCAGTCCGACAGCGCCAGGTGAAGATTCACAGAGAA 3000

QY 3001 AGAAGTTAGCGAGGAGGCTTGGACCCGCCCACTAGCAGGCTGTAGACCGCAGAGCCA 3060  
DB 3001 AGAAGTTAGCGAGGAGGCTTGGACCCGCCCACTAGCAGGCTGTAGACCGCAGAGCCA 3060  
QY 3061 AGAATTAGCCTCGCTCTGAGGAAGCGCCCTACAGGCGGTTGCTTCCGCTGGACTTTCTCT 3120  
DB 3061 AGAATTAGCCTCGCTCTGAGGAAGCGCCCTACAGGCGGTTGCTTCCGCTGGACTTTCTCT 3120  
QY 3121 AGATGCTGCTGCTGCTTACTCCAAAGTGAATCTTATATAAATCAAACTCTCTCGCACAG 3180  
DB 3121 AGATGCTGCTGCTGCTTACTCCAAAGTGAATCTTATATAAATCAAACTCTCTCGCACAG 3180  
QY 3181 GCGGAGAGCCCAATATGAGACTTGTGCTGAGCGCCCTACCTGGGGCCTTTCCACG 3240  
DB 3181 GCGGAGAGCCCAATATGAGACTTGTGCTGAGCGCCCTACCTGGGGCCTTTCCACG 3240  
QY 3241 AGCTTTGAGGGGAAAGCCATGATCTGAATATAGTATATTTCTGTAATACTCGTGAACAA 3300  
DB 3241 AGCTTTGAGGGGAAAGCCATGATCTGAATATAGTATATTTCTGTAATACTCGTGAACAA 3300  
QY 3301 ACCAAACCGGTTTTTGTCTAAGGAAAGCTAAATATGATTTTTTAAATACTATGTTTTAA 3360  
DB 3301 ACCAAACCGGTTTTTGTCTAAGGAAAGCTAAATATGATTTTTTAAATACTATGTTTTAA 3360  
QY 3361 AATACTATGTAATCTTTTTCATCTATTTAGTATATTTTATGGATGGAATAAATCTTC 3420  
DB 3361 AATACTATGTAATCTTTTTCATCTATTTAGTATATTTTATGGATGGAATAAATCTTC 3420  
QY 3421 TACTCT 3426  
DB 3421 TACTCT 3426

RESULT 6  
MUSFLK2  
LOCUS 3426 bp mRNA linear ROD 27-APR-1993  
DEFINITION Mouse flk-2 mRNA, complete cds.  
ACCESSION M64689.1  
VERSION M64689.1 GI:193327  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 3426)  
AUTHORS Matthews W., Jordan, C.T., Wiegand, G.W., Pardoll, D. and Lemischka, I.R.  
TITLE A receptor tyrosine kinase specific to hematopoietic stem and progenitor cell-enriched populations  
JOURNAL Cell 65 (7), 1143-1152 (1991)  
MEDLINE 91292518  
PUBMED 1648448  
COMMENT Original source text: Mus musculus (strain C3H/He) cDNA to mRNA.  
FEATURES  
Location/Qualifiers  
1..3426  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C3H/He"  
/db\_xref="taxon:10090"  
1..3426  
/gene="flk-2"  
31..3009  
/gene="flk-2"  
/codon\_start=1  
/protein\_id="AAA37634.1"  
/db\_xref="GI:193328"  
/translation="MRALQSRDRRLVLSVMIETVNDLPVKVLIQHNN  
GSAGKPSRYMRVSGSPEDLCQTPRQSEGTVEATVEAESGSIQLQVLPGL  
SCLWFKHSLGQPHFDLQNGIVSMALNVTETQAGEVLLHQSIRANTVLTN  
VRDTLYLRLRPYFRKMENODALLICSEGVPEPTVWVLCSSHRESCKEPAVRKE  
EKVHLEFGTDIRECCARNALGREGTKLFTIDLAQAPQSTLPQLFKVGEPLWIKKAI  
KHNHVGGLTWELEDKALEGSYFEMSTYINRIMIRILLAPVSVGRNDTGYTCSS  
KHPQSALVTILEKGFINATSQEVEIDPYEKFCFSVRPKAYPRICTWFSQSPF





QY 1921 TATGGCAATAGTAAACCGGAGTCTCAATTCAGGTGGCGTGAAGATGCTAAAGAGAA 1980  
DB 1921 TATGGCAATAGTAAACCGGAGTCTCAATTCAGGTGGCGTGAAGATGCTAAAGAGAA 1980  
QY 1981 GCTGACAGCTGTGAAGAAAGAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA 2040  
DB 1981 GCTGACAGCTGTGAAGAAAGAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA 2040  
QY 2041 CACCATGACAAATCGTGAATCTGTGGGGGATGACACTGTGAGGGCAGTGACTTG 2100  
DB 2041 CACCATGACAAATCGTGAATCTGTGGGGGATGACACTGTGAGGGCAGTGACTTG 2100  
QY 2101 ATTTTGAATATGTGTATGTGTGATGGTCACTCTCAACTACCTAAGAAAGTAAAGAGAGAAG 2160  
DB 2101 ATTTTGAATATGTGTATGTGTGATGGTCACTCTCAACTACCTAAGAAAGTAAAGAGAGAAG 2160  
QY 2161 TTTTACAGGACATGACAGAGATTTTAAAGAAATATTTTCACTTCACTTCACTTTC 2220  
DB 2161 TTTTACAGGACATGACAGAGATTTTAAAGAAATATTTTCACTTCACTTCACTTTC 2220  
QY 2221 CAGGCACATTCAAATTCAGCATGCTGTTTCAAGAGAGATTTGAACGTCTGAGCTTTGAAGAC 2400  
DB 2221 CAGGCACATTCAAATTCAGCATGCTGTTTCAAGAGAGATTTGAACGTCTGAGCTTTGAAGAC 2400  
QY 2401 CTCTTTTGTCTTGGTACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAAGTCGTGT 2460  
DB 2401 CTCTTTTGTCTTGGTACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAAGTCGTGT 2460  
QY 2461 GTTCCAGAGACCTGGCAGCCAGGATGTGTGTCACCCAGGAGGTTGAGATC 2520  
DB 2461 GTTCCAGAGACCTGGCAGCCAGGATGTGTGTCACCCAGGAGGTTGAGATC 2520  
QY 2521 TGTGACTTTGGACTGGCCGAGACATCTGAGCGACTCCAGCTACCTGTCAGGGGCAAC 2580  
DB 2521 TGTGACTTTGGACTGGCCGAGACATCTGAGCGACTCCAGCTACCTGTCAGGGGCAAC 2580  
QY 2581 GCACGGCTCCGGTGAAGTGGATGGACACCGAGAGCTTATTTGAAGGATCTACACAATC 2640  
DB 2581 GCACGGCTCCGGTGAAGTGGATGGACACCGAGAGCTTATTTGAAGGATCTACACAATC 2640  
QY 2641 AAGAGTGAGCTGTGGTCTTACGGCATCTTCTGCGAGATATTTTCACTGGGTGTAAC 2700  
DB 2641 AAGAGTGAGCTGTGGTCTTACGGCATCTTCTGCGAGATATTTTCACTGGGTGTAAC 2700  
QY 2701 CTTACCTGGCAATCTGTGCGACTAACTCTATAAATGATTCAGAGTGGATTTAA 2760  
DB 2701 CTTACCTGGCAATCTGTGCGACTAACTCTATAAATGATTCAGAGTGGATTTAA 2760  
QY 2761 ATGAGAGAGCAATCTATGTCACAGAGGATATCTTTGTAATGAAATCTCTGCTGGCT 2820  
DB 2761 ATGAGAGAGCAATCTATGTCACAGAGGATATCTTTGTAATGAAATCTCTGCTGGCT 2820  
QY 2821 TTTGACTCAAGGAAGCGGCCATCTTCCCAACCTGACTTCAATTTTGGATGTCAGTG 2880  
DB 2821 TTTGACTCAAGGAAGCGGCCATCTTCCCAACCTGACTTCAATTTTGGATGTCAGTG 2880  
QY 2881 GCAGAGCAGAAAGAGATGATCAGAAATCTCCATCCATCTACCAAAACAGCGGCCCT 2940  
DB 2881 GCAGAGCAGAAAGAGATGATCAGAAATCTCCATCCATCTACCAAAACAGCGGCCCT 2940  
QY 2941 CAGCAGAGAGCGGGCTCAGAGCCAGTGCAGCAGCGCCAGGTGAAGATTACAGAGAA 3000  
DB 2941 CAGCAGAGAGCGGGCTCAGAGCCAGTGCAGCAGCGCCAGGTGAAGATTACAGAGAA 3000  
QY 3001 AGAAGTTAGCGAGGAGGCTTTGACCCCGCCACCTAGCAGGCTGTAGACCGCAGAGCA 3060

DB 3001 AGAAGTTAGCGAGGAGGCTTGGACCCCGCCACCTAGCAGGCTGTAGACCGCAGAGCA 3060  
QY 3061 AGATTAGGCTCGCTCTGAGGAAGCGCCCTACAGGCGTGTGCTTCGTCGAGCTTTTCTCT 3120  
DB 3061 AGATTAGGCTCGCTCTGAGGAAGCGCCCTACAGGCGTGTGCTTCGTCGAGCTTTTCTCT 3120  
QY 3121 AGATGCTGTCTGCCATTACTCCAAAGTGACTTCTATAAATCAAACTCTCTCGCAGAG 3180  
DB 3121 AGATGCTGTCTGCCATTACTCCAAAGTGACTTCTATAAATCAAACTCTCTCGCAGAG 3180  
QY 3181 CGGGAGAGCCAAATTAATGAGACTTGTGTGAGCGCCGCTACCTGGGGGCTTTCCAG 3240  
DB 3181 CGGGAGAGCCAAATTAATGAGACTTGTGTGAGCGCCGCTACCTGGGGGCTTTCCAG 3240  
QY 3241 AGCTTGGGGGAAAGCCATGTATCTGAATAATATGATATTTCTGTAATACTGTAACAA 3300  
DB 3241 AGCTTGGGGGAAAGCCATGTATCTGAATAATATGATATTTCTGTAATACTGTAACAA 3300  
QY 3301 ACCAAACCCGTTTTTCTAAGGAAAGCTAAATATGATTTTAAAAATCTATGTTTTAA 3360  
DB 3301 ACCAAACCCGTTTTTCTAAGGAAAGCTAAATATGATTTTAAAAATCTATGTTTTAA 3360  
QY 3361 AATACTATGTAACCTTTTCACTATTTAGTGATATATTTATGATGAAATAAACTTTC 3420  
DB 3361 AATACTATGTAACCTTTTCACTATTTAGTGATATATTTATGATGAAATAAACTTTC 3420  
QY 3421 TACTGT 3426  
DB 3421 TACTGT 3426

## RESULT 7

144732

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches 3418;

Conservative

Mismatches

Indels

Gaps

Score 3344;

DB 6;

Length 3521;

Pred. No. 0;

0;

20;

32;

3;

1

52

61

112

121

172

181

231

240

291

300

Sequence 1 from patent US 5635388.

144732

144732

144732.1

GI:2469445

Unknown.

Unknown.

1 (bases 1 to 3521)

Bennett, B.D., Broz, S.D., Matthews, W., and Zeigler, F.C.

Agonist antibodies against the flk2/flt3 receptor and uses thereof

Patent: US 5635388-A 1 03-JUN-1997;

Location/Qualifiers

source

1..3521

/organism="unknown"

943 a 854 c 907 g 817 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches 3418;

Conservative

Mismatches

Indels

Gaps

Score 3344;

DB 6;

Length 3521;

Pred. No. 0;

0;

20;

32;

3;

1

52

61

112

121

172

181

231

240

291

300

1372	ATGTTCA	CGCTGA	TATAGA	AAAGAA	AACTCA	AGTCTG	CTAGCA	AAATGC	CTCAGC	CCAG	1431
1381	GGCTCT	GTCTCT	GTATCG	CTACCG	CTACCG	CTTAC	CTTGG	GAAG	AGTGT	TCGGA	1440
1432	GGCTCT	GTCTCT	GTATCG	CTACCG	CTACCG	CTTAC	CTTGG	GAAG	AGTGT	TCGGA	1491
1441	AAATCT	CCCAAT	TTGCA	CGGAGG	AAATCC	CAGAGG	AAATTT	GGAAAT	AAAAAG	CTTA	1500
1492	AAATCT	CCCAAT	TTGCA	CGGAGG	AAATCC	CAGAGG	AAATTT	GGAAAT	AAAAAG	CTTA	1551
1501	AAAGTGT	TTGGCC	AGTGGT	GTGAG	CTACTA	ATATAG	TGAGG	CGGGA	AAAGG	1560	
1552	AAAGTGT	TTGGCC	AGTGGT	GTGAG	CTACTA	ATATAG	TGAGG	CGGGA	AAAGG	1611	
1561	CTTCTGT	CAAAAT	CTGTG	GTATCA	ATTTCT	ATGGG	CACTCT	TTGCG	AAACAT	CTTTT	1620
1612	CTTCTGT	CAAAAT	CTGTG	GTATCA	ATTTCT	ATGGG	CACTCT	TTGCG	AAACAT	CTTTT	1671
1621	AACTCA	CCAGG	CCCTTT	CCCTTT	CATCCA	AGACAA	ATCTCT	CTATG	CGACCA	TTTGG	1680
1672	AACTCA	CCAGG	CCCTTT	CCCTTT	CATCCA	AGACAA	ATCTCT	CTATG	CGACCA	TTTGG	1731
1681	CTCTGT	CTCCCT	TTCAAT	TTGTGT	CTCAT	TTGTGT	CTCAT	TTGTGT	CTCAT	TTGTGT	1740
1732	CTCTGT	CTCCCT	TTCAAT	TTGTGT	CTCAT	TTGTGT	CTCAT	TTGTGT	CTCAT	TTGTGT	1791
1741	TTTAGG	TACAG	AGTCA	GTG	CGATG	ATCC	AGTGA	CTG	CCCTT	GGATA	1800
1792	TTTAGG	TACAG	AGTCA	GTG	CGATG	ATCC	AGTGA	CTG	CCCTT	GGATA	1851
1801	TTCTAC	GTCTG	CACTT	CAGG	CACTAT	GAAAT	TATG	ACCTT	AACTG	GGGAG	1860
1852	TTCTAC	GTCTG	CACTT	CAGG	CACTAT	GAAAT	TATG	ACCTT	AACTG	GGGAG	1911
1861	TTAGAG	TTTGG	AAAGT	CTTGG	GGTCT	GGGCT	CTGG	CGCTT	CGGAG	GGTGA	1920
1912	TTAGAG	TTTGG	AAAGT	CTTGG	GGTCT	GGGCT	CTGG	CGCTT	CGGAG	GGTGA	1971
1921	TATGGC	ATTAG	TAAAC	CGGAG	GTCT	CAAT	TTCAG	GTGG	CGTGA	AGTGA	1980
1972	TATGGC	ATTAG	TAAAC	CGGAG	GTCT	CAAT	TTCAG	GTGG	CGTGA	AGTGA	2031
1981	GCTGAC	GTCTG	TAAAG	AAAGAG	CTCT	CAAT	TTCAG	GTGG	CGTGA	AGTGA	2040
2032	GCTGAC	GTCTG	TAAAG	AAAGAG	CTCT	CAAT	TTCAG	GTGG	CGTGA	AGTGA	2091
2041	CACCAT	GACAA	CACTG	TGAA	TCTGT	GGGG	CACTG	CACTG	CACTG	CACTG	2100
2092	CACCAT	GACAA	CACTG	TGAA	TCTGT	GGGG	CACTG	CACTG	CACTG	CACTG	2151
2101	ATTTTGA	ATATG	TTGCTA	TGGTGA	CCCTCA	CTCACT	TAAG	TAAG	TAAG	TAAG	2160
2152	ATTTTGA	ATATG	TTGCTA	TGGTGA	CCCTCA	CTCACT	TAAG	TAAG	TAAG	TAAG	2211
2161	TTTCA	CAGGAC	ATGAC	GACAG	ATTTT	TAGG	AACTA	TAAT	TTCAG	TTCTT	2220
2212	TTTCA	CAGGAC	ATGAC	GACAG	ATTTT	TAGG	AACTA	TAAT	TTCAG	TTCTT	2271
2221	CAGGCA	CACTT	CAAA	TTTCC	AGCAT	GTCTG	TTTCA	CGAG	AAAGT	TCAGT	2280
2272	CAGGCA	CACTT	CAAA	TTTCC	AGCAT	GTCTG	TTTCA	CGAG	AAAGT	TCAGT	2331
2281	GATCAG	CTCT	CAGG	TTCAAT	TGGAA	TTCAAT	TTCTG	AAAGAT	GAGAT	TGAAT	2340
2332	GATCAG	CTCT	CAGG	TTCAAT	TGGAA	TTCAAT	TTCTG	AAAGAT	GAGAT	TGAAT	2391
2341	AACCA	GAGAG	GTG	CAG	AGAGG	AGGAG	AGATTT	GAAC	CGTCT	GAGAG	2400
2392	AACCA	GAGAG	GTG	CAG	AGAGG	AGGAG	AGATTT	GAAC	CGTCT	GAGAG	2451
2401	CTTCT	TTGCT	TTTGG	TAC	CAAGT	GGCC	AAAGG	CACTG	AAAT	TTCTG	2460
2452	CTTCT	TTGCT	TTTGG	TAC	CAAGT	GGCC	AAAGG	CACTG	AAAT	TTCTG	2511



172 Db |||||CTGCTGTGATCAAGTGTGTTTAAATCATAGTCAATGAGAAATATGGCTCATACGGGGAAG 231  
181 Qy |||||CCATCATCGTACCGAATCGTGGAGATCCCGAGAGACCTCCAGTGTACCCCGAGGCGC 240  
232 Db |||||CCATCATCGTACCGAATCGTGGAGATCCCGAGAGACCTCCAGTGTACCCCGAGGCGC 291  
241 Qy |||||CAGAGTGAAGGACCGGTATATGAAGCGGCACCGTGGAGGTGGCGAGTCTGGTCCATC 300  
292 Db |||||CAGAGTGAAGGACCGGTATATGAAGCGGCACCGTGGAGGTGGCGAGTCTGGSTCCATC 351  
301 Qy |||||ACCCTGCAAGTGCAGTGCACACCGAGGGACCTTCTGCTCTGGGTCTTTAAGCAC 360  
352 Db |||||ACCCTGCAAGTGCAGTGCACACCGAGGGACCTTCTGCTCTGGGTCTTTAAGCAC 411  
361 Qy |||||AGCTCCCTGGGCTGCAGCGCACTTTGATTTACAAAACAGAGGAATCGTTTCCATGGCC 420  
412 Db |||||AGCTCCCTGGGCTGCAGCGCACTTTGATTTACAAAACAGAGGAATCGTTTCCATGGCC 471  
421 Qy |||||ATCTTGAACGTGACAGAGACCCAGGCGAGGAGATACCTACTCCATATCAGAGCGAACGC 480  
472 Db |||||ATCTTGAACGTGACAGAGACCCAGGCGAGGAGATACCTACTCCATATCAGAGCGAACGC 531  
481 Qy |||||GCCAACTACACAGTACTGTTACAGTGAATGTAAAGATACACAGCTGTATGTCTAAGG 540  
532 Db |||||GCCAACTACACAGTACTGTTACAGTGAATGTAAAGATACACAGCTGTATGTCTAAGG 591  
541 Qy |||||AGACCTTACTTTAGGAAGATGGAACACAGGATGCACTGCTGCACTCCGAGGGTGT 600  
592 Db |||||AGACCTTACTTTAGGAAGATGGAACACAGGATGCACTGCTGCACTCCGAGGGTGT 651  
601 Qy |||||CGGAGCCCACTGTGAGTGGGTGCTGTCAGCTGCCAGGGAAGAGCTGTAAAGAGAA 660  
652 Db |||||CGGAGCCCACTGTGAGTGGGTGCTGTCAGCTGCCAGGGAAGAGCTGTAAAGAGAA 711  
661 Qy |||||GGCCCTGCTGTTGTGAGAAAGGAGGAAAGGTACTTCAATGAGTGTTCGGAAACAGATC 720  
712 Db |||||GGCCCTGCTGTTGTGAGAAAGGAGGAAAGGTACTTCAATGAGTGTTCGGAAACAGATC 771  
721 Qy |||||AGATGCTGTGCTAGAAATGCACTGGGCGCGAATGCAACCAAGCTTCCACATAGATCTA 780  
772 Db |||||AGATGCTGTGCTAGAAATGCACTGGGCGCGAATGCAACCAAGCTTCCACATAGATCTA 831  
781 Qy |||||AACCAGGCTCTCAGAGCACACTGCCCCAGTTATTCCTGAAAGTGGGGAAACCTTGTGG 840  
832 Db |||||AACCAGGCTCTCAGAGCACACTGCCCCAGTTATTCCTGAAAGTGGGGAAACCTTGTGG 891  
841 Qy |||||ATCAGGTGTAAAGCCATTCATGTGAACCATGGATTCGGGCTCACTGGAGCTGGAAGAC 900  
892 Db |||||ATCAGGTGTAAAGCCATTCATGTGAACCATGGATTCGGGCTCACTGGAGCTGGAAGAC 951  
901 Qy |||||AAAGCCCTGGAGGAGGCGAGCTATTTGAGATGATGACTTACTCCACAAACAGGACCATG 960  
952 Db |||||AAAGCCCTGGAGGAGGCGAGCTATTTGAGATGATGACTTACTCCACAAACAGGACCATG 1011  
961 Qy |||||ATTGCGATTTCTTTGGCTTTGTCTCTCCGTGGGAGGAAACGACACCGGATTTACACC 1020  
1012 Db |||||ATTGCGATTTCTTTGGCTTTGTCTCTCCGTGGGAGGAAACGACACCGGATTTACACC 1071  
1021 Qy |||||TGCTCTTCTCCTCAAGACACCCAGCGCAGTCAAGCTTTGGTGGACCATCTAGAAAAGGGTTT 1080  
1072 Db |||||TGCTCTTCTCCTCAAGACACCCAGCGTCAAGCTTTGGTGGACCATCTAGAAAAGGGTTT 1131  
1081 Qy |||||ATAACCGTACACAGCTCGCAAGAGATGAAATTGACCCGTAGCAAAAGTTCTGCTTC 1140  
1132 Db |||||ATAACCGTACACAGCTCGCAAGAGATGAAATTGACCCGTAGCAAAAGTTCTGCTTC 1191  
1141 Qy |||||TCAGTCAGGTTTAAAGCGTACCCAGGAATCCGATCGAGTCTTCTCAAGCCCTCA 1200  
1192 Db |||||TCAGTCAGGTTTAAAGCGTACCCAGGAATCCGATCGAGTCTTCTCAAGCCCTCA 1251  
1201 Qy |||||TTTCTTGTGACAGAGGCGCTGGAGTGGGTACAGCATATCTAAATTTTGGCATCAT 1260

1252 Db |||||TTTCTTGTGAACAGAGAGGCGCTGGAGGATGGGTACAGCATATCTTAAATTTTGGCATCAT 1311  
1261 Qy |||||AAGAACAGCCAGGAGAGTACATATTTCTATGCAAAAAATGATGACCCCGAGTTTCAACAA 1320  
1312 Db |||||AAGAACAGCCAGGAGAGTACATATTTCTATGCAAAAAATGATGACCCCGAGTTTCAACAA 1371  
1321 Qy |||||ATGTTTCAAGCTCAATATTAAGAAAGAAACCTCAAGTGTAGCAAAATGCTCAGCAGGCGAG 1380  
1372 Db |||||ATGTTTCAAGCTCAATATTAAGAAAGAAACCTCAAGTGTAGCAAAATGCTCAGCAGGCGAG 1431  
1381 Qy |||||GGGTCTGTTCCTGTATGGCTACCGGCTACCGCTCTTGGACCTTGGAGAGTGTTCGGAC 1440  
1432 Db |||||GGGTCTGTTCCTGTATGGCTACCGGCTACCGCTCTTGGACCTTGGAGAGTGTTCGGAC 1491  
1441 Qy |||||AAATCTCCCAATTTGCACGGAGGAAATCCCAAGAGGAGTTTGGAAATAAAAAAGGCTTAA 1500  
1492 Db |||||AAATCTCCCAATTTGCACGGAGGAAATCCCAAGAGGAGTTTGGAAATAAAAAAGGCTTAA 1551  
1501 Qy |||||AAAGTGTTTGGCGAGTGGGTGTGCGAGTACTCTAAATAATGAGTGAAGCGGGAAGGG 1560  
1552 Db |||||AAAGTGTTTGGCGAGTGGGTGTGCGAGTACTCTAAATAATGAGTGAAGCGGGAAGGG 1611  
1561 Qy |||||CTTCTCGTCAAAATGCTGTGCGGTACAAATTTCTATGGGACGCTTTCGGAACCATCTTTT 1620  
1612 Db |||||CTTCTCGTCAAAATGCTGTGCGGTACAAATTTCTATGGGACGCTTTCGGAACCATCTTTT 1671  
1621 Qy |||||AACTCAACGAGGCGCTTCCCTTTTCATCCAGACAAATCTCTTCTATGCAACCATTTGG 1680  
1672 Db |||||AACTCAACGAGGCGCTTCCCTTTTCATCCAGACAAATCTCTTCTATGCAACCATTTGG 1731  
1681 Qy |||||CTCTGCTCCCTTTCTATTTGTTTCTCATTTGTGATCTGCTGCAACAAATAAAAAAGCAA 1740  
1732 Db |||||CTCTGCTCCCTTTCTATTTGTTTCTCATTTGTGATCTGCTGCAACAAATAAAAAAGCAA 1791  
1741 Qy |||||TTTAGGTACGAGAGTCAAGCTGCAGATGATCCAGTGTACTGGCCCCCTGGATTAACGAGTAC 1800  
1792 Db |||||TTTAGGTACGAGAGTCAAGCTGCAGATGATCCAGTGTACTGGCCCCCTGGATTAACGAGTAC 1851  
1801 Qy |||||TTCTACGTTTGAATTTAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860  
1852 Db |||||TTCTACGTTTGAATTTAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1911  
1861 Qy |||||TTAGAGTTTGGGAGGTCCTGGGCTCTGGGCTTTGGGAGGTTGATGAAACGCCACGGCC 1920  
1912 Db |||||TTAGAGTTTGGGAGGTCCTGGGCTCTGGGCTTTGGGAGGTTGATGAAACGCCACGGCC 1971  
1921 Qy |||||TATGGCAATTAGTAAACCGGAGTCTCAAATTCAGGTGGCGGTGAAGATGCTTAAAGAGAAA 1980  
1972 Db |||||TATGGCAATTAGTAAACCGGAGTCTCAAATTCAGGTGGCGGTGAAGATGCTTAAAGAGAAA 2031  
1981 Qy |||||GCTCAAGCTGTGAAAAGAAAGCTCTCATGTGCGAGCTCAAAAATGATGATGATGATGATGATGAT 2040  
2032 Db |||||GCTCAAGCTGTGAAAAGAAAGCTCTCATGTGCGAGCTCAAAAATGATGATGATGATGATGATGAT 2091  
2041 Qy |||||CACATGACACATCTGTGAATCTGCTGGGCGATGACACTGTGAGGGCGAGTACTTTG 2100  
2092 Db |||||CACATGACACATCTGTGAATCTGCTGGGCGATGACACTGTGAGGGCGAGTACTTTG 2151  
2101 Qy |||||ATTTTGTGAATTTGTTGCTATGTTGACCTCTCAACTACCTAAGAGTAAAGAGAGAGAG 2160  
2152 Db |||||ATTTTGTGAATTTGTTGCTATGTTGACCTCTCAACTACCTAAGAGTAAAGAGAGAGAG 2211  
2161 Qy |||||TTTCAAGGACATGGACAGAGATTTTAAAGGAACATAATTTTCAGTCTTTACCCCTACTTTC 2220  
2212 Db |||||TTTCAAGGACATGGACAGAGATTTTAAAGGAACATAATTTTCAGTCTTTACCCCTACTTTC 2271  
2221 Qy |||||CAGGACATTTCAATTTCCAGCATGCTGTTTCCAGAGAGTTTCAAGTACCCCGCCCTTG 2280  
2272 Db |||||CAGGACATTTCAATTTCCAGCATGCTGTTTCCAGAGAGTTTCAAGTACCCCGCCCTTG 2331  
2281 Qy |||||GATCAGCTCTCAGGCTTCAATGGGAAATTCAAATTCATTTCTGAAGATGAGATTTGAATGA 2340  
2332 Db |||||GATCAGCTCTCAGGCTTCAATGGGAAATTCAAATTCATTTCTGAAGATGAGATTTGAATGA 2391

```
QY 2341 AACCAAGAGGCTGGCGCAGAGAGAGAGAGATTGTAACGTGCTGACGTTTGAAGAC 2400
Db 2392 AACCAAGAGGCTGGCGCAGAGAGAGAGAGATTGTAACGTGCTGACGTTTGAAGAC 2451
QY 2401 CTCCTTTTCTTTCGCTACCAAGTGGCCAAAGGATGGAATTCCTCGAGTTCAGTCTGT 2460
Db 2452 CTCCTTTTCTTTCGCTACCAAGTGGCCAAAGGATGGAATTCCTCGAGTTCAGTCTGT 2511
QY 2461 GTCCACAGAGACTGGCAGCCAGGAATGTGTGTCACCCAGCGGAGGTGTGAAGATC 2520
Db 2512 GTCCACAGAGACTGGCAGCCAGGAATGTGTGTCACCCAGCGGAGGTGTGAAGATC 2571
QY 2521 TGTGACTTTGGACTGGCCCGAGACATCTCTGAGGACTCCAGCTACGTCTGTCAGGGCAAC 2580
Db 2572 TGTGACTTTGGACTGGCCCGAGACATCTCTGAGGACTCCAGCTACGTCTGTCAGGGCAAC 2631
QY 2581 GCACGGCTCCCGGTGAAGTGGATGGCCACCGAGAGCTTATTTGAAGGATCTACACAATC 2640
Db 2632 GCACGGCTCCCGGTGAAGTGGATGGCCACCGAGAGCTTATTTGAAGGATCTACACAATC 2691
QY 2641 AAGAGTGAGCTGTGGTCTACGGCATCTCTCTGCGAGATATTTTCACTGGGTGTGAAC 2700
Db 2692 AAGAGTGAGCTGTGGTCTACGGCATCTCTCTGCGAGATATTTTCACTGGGTGTGAAC 2751
QY 2701 CCTTACCTGGCATCTCTCTGAGCTAATCTTATATAAATGATTCAGATGGAATTTAA 2760
Db 2752 CCTTACCTGGCATCTCTCTGAGCTAATCTTATATAAATGATTCAGATGGAATTTAA 2811
QY 2761 ATGGAGCAGCAATCTATGCCACAGAGGATATCTTTGTAATGCAATCTCTGCGGT 2820
Db 2812 ATGGAGCAGCAATCTATGCCACAGAGGATATCTTTGTAATGCAATCTCTGCGGT 2871
QY 2821 TTTGACTCAAGGAAGCGGCATCTCTCCCAACCTGACTTCAATTTTAGATGTGAGCTG 2880
Db 2872 TTTGACTCAAGGAAGCGGCATCTCTCCCAACCTGACTTCAATTTTAGATGTGAGCTG 2931
QY 2881 GCAGAGGAGAGAAGC-----ATGTATCAAGAAATCATCATCC 2917
Db 2932 GCAGAGGAGAGAAGAGCGATGTATCAGAAATGCGGTGGCAAGCTCCAGAAATCATCATCC 2991
QY 2918 ATCTACCAAAACAGCGGCCCTCAGCAGAGAGGCGGCTCAGAGCC-CAGTGGCCACAG 2976
Db 2992 ATCTACCAAAACAGCGGCCCTCAGCAGAGAGGCGGCTCAGAGCCCGCCATCGCCACAG 3051
QY 2977 CGCCAGGTGAAGATTACAGAGAAAGATTTAGCGAGAGGCTTTGGACCCCGCCACCT 3036
Db 3052 GCCCAGGTGAAGATTACAGAGAAAGATTTAGCGAGAGGCTTTGGACCCCGCCACCT 3111
QY 3037 AGCAGGCTGTAGACCGCAGAGCGCAAGATTAGCTCGCTCTGAGGAAGCGCCCTACAGCG 3096
Db 3112 AGCAGGCTGTAGACCGCAGAGCGCAAGATTAGCTCGCTCTGAGGAAGCGCCCTACAGCG 3171
QY 3097 CGTTGCTTCGCTGGACTTTTCTCTAGATGCTGCTGCCATTAATCTCCTCAAGTACTTCTAT 3156
Db 3172 CGTTGCTTCGCTGGACTTTTCTCTAGATGCTGCTGCCATTAATCTCCTCAAGTACTTCTAT 3231
QY 3157 AAAATCAAACTCTCTCGCAGCGGAGAGCAATAATAGACTTTGTTGGTGAGGCC 3216
Db 3232 AAAATCAAACTCTCTCGCAGCGGAGAGCAATAATAGACTTTGTTGGTGAGGCC 3291
QY 3217 GCCTACCTGGGGCCCTTT-----CCAGAGCTTGGGGGAAAGCCATGTATCTGAA 3268
Db 3292 GCCTACCTGGGGCCCTTTCCAGGCCCCCGAGCTTGGGGGAAAGCCATGTATCTGAA 3351
QY 3269 ATATAGTATATCTTTGTAATACGTGAAACAAACCAACCGCTTTTGTCTAAGGGAAG 3328
Db 3352 ATATAGTATATCTTTGTAATACGTGAAACAAACCAACCGCTTTTGTCTAAGGGAAG 3411
QY 3329 CTAATATGATTTTTTAAATCTATGTTTTTAAATATCTATGTAATCTTTTCAATCTATTTA 3388
Db 3412 CTAATATGATTTTTTAAATCTATGTTTTTAAATATCTATGTAATCTTTTCAATCTATTTA 3471
```

```
QY 3389 GTGATATATTTTATGATGGAATAAACTTTCTACTGTAAAAA 3438
Db 3472 GTGATATATTTTATGATGGAATAAACTTTCTACTGTAAAAA 3521

RESULT 9
LOCUS AX695520 2979 bp DNA linear PAT 31-MAR-2003.
DEFINITION Sequence 1147 from Patent WO03008583.
ACCESSION AX695520
VERSION AX695520.1 GI:29418672
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Morris, D.W. and Engelhard, E.K.
TITLE Novel compositions and methods for cancer
JOURNAL Patent: WO 03008583-A 1147 30-JAN-2003;
Sagres Discovery (US)
FEATURES
source
1..2979
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
BASE COUNT 803 a 714 c 775 g 687 t
ORIGIN

Query Match 86.3%; Score 2979; DB 6; Length 2979;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2979; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 ATCGGGCGTTGGCCAGCGAGCAGCGGGCTGCTGCTCTTTGTCAGTA 90
Db 1 ATCGGGCGTTGGCCAGCGAGCAGCGGGCTGCTGCTCTTTGTCAGTA 60
QY 91 ATGATTTCTTGAGACCGTTTACAAACCAAGACCTGCTGCTGATCAAGTGTGTTAATCACT 150
Db 61 ATGATTTCTTGAGACCGTTTACAAACCAAGACCTGCTGCTGATCAAGTGTGTTAATCACT 120
QY 151 CATGAGAAATATGGCTCATACGCGGAAAGCCATCATCGTACCAGTATGTCGAGATCC 210
Db 121 CATGAGAAATATGGCTCATACGCGGAAAGCCATCATCGTACCAGTATGTCGAGATCC 180
QY 211 CCAGAGACCTCTCAGTGTACCCCGAGGCGCAGAGTGAAGGACGGTATATGAAGCGGCC 270
Db 181 CCAGAGACCTCTCAGTGTACCCCGAGGCGCAGAGTGAAGGACGGTATATGAAGCGGCC 240
QY 271 ACCGTGGAGTGGCCGAGTCTGGGTCCATCACCTTGCAGTGCAGCTGCCACCCAGGG 330
Db 241 ACCGTGGAGTGGCCGAGTCTGGGTCCATCACCTTGCAGTGCAGCTGCCACCCAGGG 300
QY 331 GACCTTTCTGCTCTCTGGGTCTTTAAGCAGACAGCTCCCTGGGCTGCCAGCCGACATTTGAT 390
Db 301 GACCTTTCTGCTCTCTGGGTCTTTAAGCAGACAGCTCCCTGGGCTGCCAGCCGACATTTGAT 360
QY 391 TTACAAACAGAGGAATCGTTTTCATGGCCATCTTGAACGTGACAGAGACCAGGACGGA 450
Db 361 TTACAAACAGAGGAATCGTTTTCATGGCCATCTTGAACGTGACAGAGACCAGGACGGA 420
QY 451 GAATPACCTACTCCATATTCAGAGGAGCGGCAACTACACAGTACTGTTTCAAGTGAAT 510
Db 421 GAATPACCTACTCCATATTCAGAGGAGCGGCAACTACACAGTACTGTTTCAAGTGAAT 480
QY 511 GTAAGAGATACACAGCTGTATGTCTAAGAGACCTTACTTTAGGAAGATGGAAGAACAG 570
Db 481 GTAAGAGATACACAGCTGTATGTCTAAGAGACCTTACTTTAGGAAGATGGAAGAACAG 540
QY 571 GATGCACTGCTCTGCACTCTCCGAGGGTGTTCGAGGCCCACTGTGGAGTGGGTCTCTGC 630
Db 541 GATGCACTGCTCTGCACTCTCCGAGGGTGTTCGAGGCCCACTGTGGAGTGGGTCTCTGC 600
```



QY	631	AGCTCCCA	CAGGAAAGCTGTAAGAGAGAGCCCTGCTGTTGTCAGAAAGGAGGAAAAG	690	Db	1681	GTGTTGAT	CTGCCACAAATACAAAAGCAATTTAGGTACGAGAGTCAGCTGCAGATGATC	1740
Db	601	AGCTCCCA	CAGGAAAGCTGTAAGAGAGAGCCCTGCTGTTGTCAGAAAGGAGGAAAAG	660	QY	1771	CAGGTGAT	CGGCCCTTGGATTAACGAGTACTTCTACGTTGACTTCAGGGACTATGAATAT	1830
QY	691	GTACTTCA	TGAGTTGTCGGAACAGACATCAGATGCTGTAGAAATGCACTGGGCGGC	750	Db	1741	CAGGTGAT	CGGCCCTTGGATTAACGAGTACTTCTACGTTGACTTCAGGGACTATGAATAT	1800
Db	661	GTACTTCA	TGAGTTGTCGGAACAGACATCAGATGCTGTAGAAATGCACTGGGCGGC	720	QY	1831	GACCTTAA	GTGGGAGTTCCCGAGAGAGAACTTAGAGTTTGGGAGGTCTCTGGGGTCTGGC	1890
QY	751	GAATGACA	CAAGCTGTTACCATAGATCTAAACAGGCTCTCTCAGAGCACTGCCCCAG	810	Db	1801	GACCTTAA	GTGGGAGTTCCCGAGAGAGAACTTAGAGTTTGGGAGGTCTCTGGGGTCTGGC	1860
Db	721	GAATGACA	CAAGCTGTTACCATAGATCTAAACAGGCTCTCTCAGAGCACTGCCCCAG	780	QY	1891	GCCTTTGG	AGGGGTGATGAACGCCACCGCTATATGGCAATTAGTAAACCGGGAGTCTCAATT	1950
QY	811	TTATTTCT	GAAAGTGGGGAAACCTTGTGATCAGGTGTAAGCCATCCATGTGAACCAT	870	Db	1861	GCCTTTGG	AGGGGTGATGAACGCCACCGCTATATGGCAATTAGTAAACCGGGAGTCTCAATT	1920
Db	781	TTATTTCT	GAAAGTGGGGAAACCTTGTGATCAGGTGTAAGCCATCCATGTGAACCAT	840	QY	1951	CAGGTGCG	GTGAAGATGCTTAAAGAGAAAGCTGACAGCTGTGTGAAAAGAAAGCTCTCATG	2010
QY	871	GGATTCCG	GGCTCACCCTGGGAGCTGGAAGACAAAGCCCTGGAGGAGGCGACTCTTTGAG	930	Db	1921	CAGGTGCG	GTGAAGATGCTTAAAGAGAAAGCTGACAGCTGTGTGAAAAGAAAGCTCTCATG	1980
Db	841	GGATTCCG	GGCTCACCCTGGGAGCTGGAAGACAAAGCCCTGGAGGAGGCGACTCTTTGAG	900	QY	2011	TCGGAGCT	CAAAATGATGACCCCTGGGACACCATGACAACTCGTGAATCTGCTGGG	2070
QY	931	ATGAGTAC	CTTACTCCAAACAGGACCATGATTGGGATCTCTTTGGGCTTTGTGCTTCC	990	Db	1981	TCGGAGCT	CAAAATGATGACCCCTGGGACACCATGACAACTCGTGAATCTGCTGGG	2040
Db	901	ATGAGTAC	CTTACTCCAAACAGGACCATGATTGGGATCTCTTTGGGCTTTGTGCTTCC	960	QY	2071	GCATGCAC	ACTGTGAGGCGCAGTGACTTGTGATTTTGAATTTTGTGCTATGTTGACCTC	2130
QY	991	GTGGGAAG	GAGACACCGGATATTACACCTGCTTCTCAAGCACCCCGCCAGCTCA	1050	Db	2041	GCATGCAC	ACTGTGAGGCGCAGTGACTTGTGATTTTGAATTTTGTGCTATGTTGACCTC	2100
Db	961	GTGGGAAG	GAGACACCGGATATTACACCTGCTTCTCAAGCACCCCGCCAGCTCA	1020	QY	2131	CTCAACTAC	CTTAAGATTAAGAGAGAAAGTTTCAAGGACATGGAACAGAGATTTTAAAG	2190
QY	1051	GCCTTGGT	GACCATCTAGAAAAGGGTTTATAACGCTACCAAGCTCGCAAGAGAGTAT	1110	Db	2101	CTCAACTAC	CTTAAGATTAAGAGAGAAAGTTTCAAGGACATGGAACAGAGATTTTAAAG	2160
Db	1021	GCCTTGGT	GACCATCTAGAAAAGGGTTTATAACGCTACCAAGCTCGCAAGAGAGTAT	1080	QY	2191	GAACATTA	TTTTCAGTTCTTACCTTCCAGGCACATTCAAATTCAGCATGCTCGT	2250
QY	1111	GAATTTGA	CCCGTACGAAAAGTTCTGCTTCTCAGTCAGGTTTAAAGGGTACCCAGCAATC	1170	Db	2161	GAACATTA	TTTTCAGTTCTTACCTTCCAGGCACATTCAAATTCAGCATGCTCGT	2220
Db	1081	GAATTTGA	CCCGTACGAAAAGTTCTGCTTCTCAGTCAGGTTTAAAGGGTACCCAGCAATC	1140	QY	2251	TCAGGAGA	GTTCAGTTTACACCCGCCCTTGGATCAGCTCTCAGGGTTCATTTGGGAATTTCA	2310
QY	1171	CGATGCA	GTGGATCTTCTCAAGCCTCATTTCTTGTGGAACAGAGAGGCGCTGGAGAT	1230	Db	2221	TCAGGAGA	GTTCAGTTTACACCCGCCCTTGGATCAGCTCTCAGGGTTCATTTGGGAATTTCA	2280
Db	1141	CGATGCA	GTGGATCTTCTCAAGCCTCATTTCTTGTGGAACAGAGAGGCGCTGGAGAT	1200	QY	2311	ATTCATTT	CTGAAGATGAGATTGAATATGAATAAGAGAGGCTGGCAGAGAGAGAGAG	2370
QY	1231	GGGTACAG	CATATCTAAATTTTGGATCATAGAACCAAGCCAGGAGAGTACATATTTCTAT	1290	Db	2281	ATTCATTT	CTGAAGATGAGATTGAATATGAATAAGAGAGGCTGGCAGAGAGAGAGAG	2340
Db	1201	GGGTACAG	CATATCTAAATTTTGGATCATAGAACCAAGCCAGGAGAGTACATATTTCTAT	1260	QY	2371	GAAGATTG	AAAGTGTGTCGACCTCTTTGGAAGACCTCTTTGCTTTGCTACCAAGTGGCCAA	2430
QY	1291	GCAGAAAT	GATGATGCGCCAGTTCACCAAAATGTTTCAGCTGAAATATAAGAAAGAAACCT	1350	Db	2341	GAAGATTG	AAAGTGTGTCGACCTCTTTGGAAGACCTCTTTGCTTTGCTACCAAGTGGCCAA	2400
Db	1261	GCAGAAAT	GATGATGCGCCAGTTCACCAAAATGTTTCAGCTGAAATATAAGAAAGAAACCT	1320	QY	2431	GGCATGAA	TTTCTTGGAGTTCAAGTCTGCTCCACAGAGACCTGGCAGCCAGGAAATGTG	2490
QY	1351	CAAGTGCT	AGCAATGCTCAGCCAGCGCGCTCTGTTCTCTGATGGCTACCCGCTA	1410	Db	2401	GGCATGAA	TTTCTTGGAGTTCAAGTCTGCTCCACAGAGACCTGGCAGCCAGGAAATGTG	2460
Db	1321	CAAGTGCT	AGCAATGCTCAGCCAGCGCGCTCTGTTCTCTGATGGCTACCCGCTA	1380	QY	2491	TTGGTCA	CCCCACGGGAAGTGGTGAAGATCTGTGACTTTGGAATGGCCCGGAGACATCCTG	2550
QY	1411	CCCTTTGG	ACCTGGAAGATGTTTCGGAACAAATCTCCCAATTCGACGGAGGAAATCCCA	1470	Db	2461	TTGGTCA	CCCCACGGGAAGTGGTGAAGATCTGTGACTTTGGACTGGCCCGGAGACATCCTG	2520
Db	1381	CCCTTTGG	ACCTGGAAGATGTTTCGGAACAAATCTCCCAATTCGACGGAGGAAATCCCA	1440	QY	2551	AGCGACTC	CCAGCTACGTCAGGGGCAACGCACGGCTGCGGGTGAAGTGAATGGCAACC	2610
QY	1471	GAAGGAGT	TTTGAATATAAGGCTTAACAGAAAGTTTGGCCAGTGGGTGTCGAGCAGT	1530	Db	2521	AGCGACTC	CCAGCTACGTCAGGGGCAACGCACGGCTGCGGGTGAAGTGAATGGCAACC	2580
Db	1441	GAAGGAGT	TTTGAATATAAGGCTTAACAGAAAGTTTGGCCAGTGGGTGTCGAGCAGT	1500	QY	2611	GAGAGCTT	ATTTTGAAGGATCTACACATCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT	2670
QY	1531	ACTCTAA	TATAGTGGGCGGGAAGGGCTTCTGGTCAAAATGCTGTGCGTACAAATTTCT	1590	Db	2581	GAGAGCTT	ATTTTGAAGGATCTACACATCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT	2640
Db	1501	ACTCTAA	TATAGTGGGCGGGAAGGGCTTCTGGTCAAAATGCTGTGCGTACAAATTTCT	1560	QY	2671	CTCTGGG	AGATATTTTCACTGGGTGTAACCTTACCTCGGCAATTCCTGTGACGCTAAC	2730
QY	1591	ATGGGCA	CGTCTTGGAAACCATCTTTTAAACTCACCGAGGCCCTTCCCTTTATCCAA	1650	Db	2641	CTCTGGG	AGATATTTTCACTGGGTGTAACCTTACCTCGGCAATTCCTGTGACGCTAAC	2700
Db	1561	ATGGGCA	CGTCTTGGAAACCATCTTTTAAACTCACCGAGGCCCTTCCCTTTATCCAA	1620	QY	2731	TTCTATAA	ACTGATTCAGAGTGGATTTAAATATGGAGCAGCAATTCCTATGCAAGAGGG	2790
QY	1651	GACAACTC	CTCTTATGACGACCATTTGGGCTCTGCTCCCTTCTCATTTGTTTCTCAT	1710	Db	2701	TTCTATAA	ACTGATTCAGAGTGGATTTAAATATGGAGCAGCAATTCCTATGCAAGAGGG	2760
Db	1621	GACAACTC	CTCTTATGACGACCATTTGGGCTCTGCTCCCTTCTCATTTGTTTCTCAT	1680	QY	2791	ATATAC	TTTGTAAATGCAATCTCTGCTGGCTTTTGAATCAAGGAAGCGGCATCTCTCC	2850
QY	1711	GTGTTGAT	CTGCCACAAATACAAAAGCAATTTAGGTACGAGAGTCAGCTGCAGATGATC	1770					

Db 2761 ATATACTTGTAAATGCAATCTGCTGGGCTTTTGAATCAAGGAAGCGGCATCTTCCCC 2820  
QY 2851 AACCTGACTTCAATTTTAGATGTCAGTGGGAGAGGAGAGAGATGATCAGAA 2910  
Db 2821 AACCTGACTTCAATTTTAGATGTCAGTGGGAGAGGAGAGAGATGATCAGAA 2880  
QY 2911 TCCATCCATCTACCAAAACAGGGGGCCCTCAGCAGAGAGGGGGCTCAGAGCCAGTCG 2970  
Db 2881 TCCATCCATCTACCAAAACAGGGGGCCCTCAGCAGAGAGGGGGCTCAGAGCCAGTCG 2940  
QY 2971 CCACAGCGCCAGGTGAAGATTACAGAGAAAGAAAGTTAG 3009  
Db 2941 CCACAGCGCCAGGTGAAGATTACAGAGAAAGAAAGTTAG 2979

RESULT 10  
LOCUS AR005212 3501 bp DNA linear PAT 04-DEC-1998  
DEFINITION Sequence 3 from patent US 5747651.  
ACCESSION AR005212  
VERSION AR005212.1 GI:3966091  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 3501)  
AUTHORS Lemischka, I.R.  
TITLE Antibodies against tyrosine kinase receptor flk-1  
JOURNAL Patent: US 5747651-A 3 05-MAY-1998;  
FEATURES  
Location/Qualifiers  
1..3501  
/organism="unknown"  
BASE COUNT 1068 a 709 c 784 g 940 t  
ORIGIN

Query Match 64.5%; Score 2227.2; DB 6; Length 3501;  
Best Local Similarity 79.6%; Pred. No. 0;  
Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;

QY 8 GGCTACCGCGGCTCCGAGGCGCATGCGGGCGTTGGCGAGCGCGACCGCGCGCTGC 67  
Db 35 GGGGACCCCGGCTCCGAGGCGCATGCGGGCGTTGGC---GGCGACGGGSCACCGTGC 91  
QY 68 TGCTGCTTGTGTTTGTGTAAGTATCTTGACACCGTTACAAACAGACCTGCCTG 127  
Db 92 CGCTGCTGCTGTTTCTGCAATGATATTTGGGACTATTACAAATCAAGATCTGCGTG 151  
QY 128 TGATCAAGTGTGTTTAAATCAATGATGAGAAATGATGATGATGATGATGATGATGAT 187  
Db 152 TGATCAAGTGTGTTTAAATCAATGATGAGAAATGATGATGATGATGATGATGATGAT 211  
QY 188 CGTACCGGAATGGTGGAGGATCCCGAGAGACCTCCAGTGTACCCCGAGGCGCCAGAGT 247  
Db 212 CATATCCATGGTATCAAAATCCCGGAGACCTCGGGTGGTGGAGACCCAGACT 271  
QY 248 AAGGACGGTATGAGCGGCGACCGTGGAGGTGGCGAGTGGGGTCCATCAACCTGC 307  
Db 272 CAGGACAGTGTACGAAAGTCCCGTGTGGAAGTGGATGATGATGATGATGATGATGATG 331  
QY 308 AAGTGCAGTCCGACCCCGGAGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 367  
Db 332 AAGTGCAGTCCGACCCCGGAGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 391  
QY 368 TGGGCTGCCAGCGGACCTTTGATTTACAAACAGAGGAATCGTTTCCATGGCCATCTTGA 427  
Db 392 TGAATGCCAGCCACATTTTGAATTTACAAACAGAGGAGTGTGTTCCATGTCATTTGA 451  
QY 428 ACGTGACAGAGACCCAGGAGGAAATACCTACTCCATATTTCAGAGCGAAACGCGCAACT 487  
Db 452 AAATGACAGAAACCCAGCTGGAGAAATACCTACTTTTATTTCAGAGTGAAGCTACCAAT 511  
QY 488 ACACAGTACTGTTTACAGTGAATGTAAGAGATACACAGCTGATGCTGAAGGACCTT 547

Db 512 ACACAAATTTGTTTACAGTGAGTAAAGAAATACCCCTGCTTTTACACATTAAGAAGACCTT 571  
QY 548 ACTTTAGGAAGATGGAACACAGGATGCACTGCTCTGCATCTCCGAGGGTGTTCGCGAGC 607  
Db 572 ACTTTAGAAAATGGAACACAGGACGCGCTTGTCTGCATATCTGAGAGCGTTCAGAGC 631  
QY 608 CCACCTGTGAGTGGTGTCTCTGAGCTCCACAGGAAAAGCTGTAAGAAGAAAGCGCCCTG 667  
Db 632 CGATCGTGAATGGGTGCTTTGCGATTACAGGGGGGAAAGCTGTAAGAAGAAAGTCCAG 691  
QY 668 CTGTTGTCAAGAAAGAGGAAAAGTACTTTCATGAGTTGTTCCGAAACAGACATCAGATGCT 727  
Db 692 CTGTTGTTAAAGAGGAAAAGTGTCTTCAAGAAATTAATGGGACGGAATGAAGTGTCT 751  
QY 728 GTGCTAGAAATGCACTGGGCGCGCAATGCACAAAGCTGTTCCACATAGATCTAAACAGG 787  
Db 752 GTGCCAGAAATGAACCTGGCAGGGAATGCACAGGCTGTTCAATATAGATCTAAATCAAA 811  
QY 788 CTCCTCAGAGCACACTGCCCAAGTTATTCCTGAAAGTGGGGGAAACCTTTGTGATCAGGT 847  
Db 812 CTCCTCAGACCAATGTCACCAATTTATTTCTTAAAGTAGGGGAACCTTTATGATAAGGT 871  
QY 848 GTAAGGCCATCCATGTAACCATGATTCGGGCTCACCTGGGAGCTGGGAAGCAAAAGCCC 907  
Db 872 GCMAAGCTGTTTCATGTGAACCAATGATTCGGGCTCACCTGGGAATTTAGAAAACAAAGCAC 931  
QY 908 TGGAGGAGGCGAGCTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATGATTCGGA 967  
Db 932 TCGAGGAGGCGCAACTACTTTGAGATGAGTACCTATTCAACAAACAGCACTATGATACGGA 991  
QY 968 TTCTCTGGCCTTTGTGCTTCCGTTGGGAAGAAACGACACCGGATATTACACCTGCTCTT 1027  
Db 992 TTCTGTTGCTTTTGTATCATCAGTGGCAAGAAACGACACCGGATATTACACCTGTTCT 1051  
QY 1028 CCTCAAGCACCCAGCCAGCTCAGCTGGTGGTGAACCATCTCTAGAAAAGGGTTTATAACG 1087  
Db 1052 CTTCAAGCATCCCAAGTCAATCAGCTTTGGTTTACCATCTGTAGAAAAGGGATTTATAAATG 1111  
QY 1088 CTACAGCTCCCAAGAGATGATAATGACCGCTACGAAAGTTCTGCTTCTCAGTCA 1147  
Db 1112 CTACCAATTCAGTGAAGATTATGAAATGACCAATATGAAGATTTGTTTTCTGTCA 1171  
QY 1148 GGTTTAAAGCGTACCCAGAAATCCGATGACGTTCTCTCAAGCTCATTTCTCTT 1207  
Db 1172 GGTTTAAAGCGTACCCAAATCAGATGATGATGATGATGATGATGATGATGATGATGATG 1231  
QY 1208 GTGAACAGAGAGGCTGGAGATGGGTACAGCATATCTAAATTTTGGCATATAAGACA 1267  
Db 1232 GTGAGCAAAAGGGTCTTGATAACGGATACAGCATATCCAAAGTTTGGCAATCATAGCAC 1291  
QY 1268 AGCCAGGAGATACATATTTCTATCAGAAAATGATGACGCCAGTTTCCAAAATGTTCA 1327  
Db 1292 AGCCAGGAGATATATTTCTATCAGAAAATGATGATGATGATGATGATGATGATGATG 1351  
QY 1328 CGCTGAATATAAGAAAGAAACCTCAAGTGTCTAGCAAAATGCTCTCAGCCAGCGGCTCT 1387  
Db 1352 CGCTGAATATAAGAAAGAAACCTCAAGTGTCTCAGAGCATATCGGCAAGTCAAGGCTCT 1411  
QY 1388 GTTCTCTGATGGTACCCGCTACCTTCTGGAATCTGGAAGAGTGTTCGGAACAATCTC 1447  
Db 1412 GTTCTCTGGAATGGAATACCTTACATCTTGGACCTGGAAGAGTGTTCAGACAAGTCTC 1471  
QY 1448 CCAATTCAGCGAGGAAATCCCAAGAGGTTTGGAAATAAAGGCTTAACAGAAAGTGT 1507  
Db 1472 CCACTGCAGAGAGATCAGAGAGAGTCTGGAATAGAAAGGCTTAACAGAAAGTGT 1531  
QY 1508 TTGCGCAGTGGGTCTCGAGCAGTACTCTAAATATGATGAGGGCGGGGAAAGGGCTTCTGG 1567  
Db 1532 TTGACAGTGGGTCTCGAGCAGTACTCTAAACATGATGAGGAGCCATAAAGGGTTCCTGG 1591  
QY 1568 TCAATGCTGTGCTACAAATCTATGAGGACGCTTTCGGAACCATCTTTTTAACTCAC 1627  
Db 1592 TCAAGTCTGTGCATACAAATTCCTTGGCACATCTTGTGAGACGATCTCTTTTAACTCTC 1651



BASE COUNT	1068 a	709 c	784 g	940 t	
ORIGIN	/organism="unknown"				
Query Match	64.5%; Score 2227.2; DB 6; Length 3501;				
Best Local Similarity	79.6%; Pred. No. 0;				
Matches 2775; Conservative	0; Mismatches 653; Indels 56; Gaps 10;				
Qy	8	GGCTACCGCGCTCCGAGGCGCATGCGGGCTTGGCGCAGCGCAGCGCGCGCTGC	67		
Db	35	GGGACCCCGGCTCCGAGGCGCATGCGGGCTTGGC---GGCGACCGGGCACCGTGC	91		
Qy	68	TGCTGCTTGTGTTTGTTCAGTAATGATTTCTTGACACCGTTTACAAACCAAGACCTGCCTG	127		
Db	92	CGTGTCTGTTGTTTCTGCAATGATATTGGGACTATTACAATCAAGATCTGCGCTG	151		
Qy	128	TGATCAAGTGTGTTTAAATCAGTCATGAGAAATGCGTCTACAGCGGGAAGCCATCAT	187		
Db	152	TGATCAAGTGTGTTTAAATCAATCAATAAGAAATGATTCATCAGTGGGAAGTCATCAT	211		
Qy	188	CGTACCGAATGTCGAGATCCCGAAGACCTCCAGTGTACCCCGAGGGCCGAGGTG	247		
Db	212	CATATCCCATGGTATCAGATCCCGGAAGACCTCGGTGTGCGTTGAGACCCGAGACT	271		
Qy	248	AAGGACGGTATATGAAGCGCCACCGTGGAGGTGGCGAGTCTGGTCCATCACCTGC	307		
Db	272	CAGGACAGTGTACGAGCTGCGGTGTGGAAGTGGATGTATCTGCTCCATCACACTGC	331		
Qy	308	AAGTGCAGCTCCGACCCGAGGACCTTCTGCTGCTGCGTCTTAAAGCACAGCTCCC	367		
Db	332	AAGTGTGCTCGATGCGCCGAGGAAATTTCTGCTCTGGGTCTTTAAGCACAGCTCCC	391		
Qy	368	TGGCTGCCAGCCAGCTTTGATTTTAAACAGAGGAATCGTTTCCATGGCCATCTTGA	427		
Db	392	TGAATGCCAGCCACATTTTGAATTTACAAACAGAGGAGTGTTCATGCTCATTTGA	451		
Qy	428	ACGTGACAGACACCGAGGAGATACCTTCTCATATTACAGAGCAACCGCCCACT	487		
Db	452	AAATGACAGAAACCCAGCTGGAGATACCTTCTTTTATTCAGATGAAGTACCATT	511		
Qy	488	ACACAGTACTGTTTCAAGTGAATGTAAGAGATACACAGCTGTATGTGCTAAGAGACCTT	547		
Db	512	ACACATATTGTTTACAGTGAATATAAGAAATACCTGCTTTACATTAAGAGACCTT	571		
Qy	548	ACTTTAGGAATGGAACACAGATGACCTGCTCTGATCTCCGAGGCTGTTCCGAGC	607		
Db	572	ACTTTAGAAAAATGGAACACAGACCGCTGCTGCTGATATCTGAGAGCGTTCCAGAGC	631		
Qy	608	CCACTGTGAGTGGTCTCTGACGCTCCACAGGGAAGCTGTAAGAGAGAGCCCTG	667		
Db	632	CGATCTGGAATGGTCTTTCGATTTCAAGGGGAAGCTGTAAGAGAGAGAGTCCAG	691		
Qy	668	CTGTGTGACAGAGGAGAAAGTACTTCTATGATGTTGTCGGAACAGACATCAGATGCT	727		
Db	692	CTGTGTGTAAGAGGAGAAAGTGTCTTATGATATTATTGGACGACATTAAGTGTCT	751		
Qy	728	GTGCTAGAAATGACCTGGCGGGAATGACCAAGCTGTTTCAATAGATCTAAACGAG	787		
Db	752	GTGCCAGAAATGAATCGGGCAGGGAATGCAACAGGCTGTTTCAATATAGATCTAAATCAA	811		
Qy	788	CTCCTCAGACACACTGCCCCAGTTATCTCTGAAAGTGGGGAAACCTCTGATCAGGT	847		
Db	812	CTCCTCAGACCATTTGCCCAATTTATTTCTTAAAGTAGGGAAACCTTTATGATAGGT	871		
Qy	848	GTAAGGCCATCCATGTGAACCATAGATTCGGGCTCACCTGGGAGCTGGAAGACAAAGCCC	907		
Db	872	GCAAGCTGTTTATGTGAACCATGATTCGGGCTCACCTGGGATTTAGAAAACAAAGCAC	931		
Qy	908	TGAGAGGGCAGCTACTTTGAGATGATGATCTTCACTCAACAAAGGACCATGATTCGGA	967		
Db	932	TCAGAGGGGCAACTACTTTGATGATGATGATGATGATGATGATGATGATGATGATGAT	991		
Qy	968	TTCTCTTGGCTTTGTGCTCTCTCGTGGGAGGAAACGACCGGATATTACCTGCTCTT	1027		

Db	992	TTCTGTTTCTTTTGTATCATAGTGGCAAGAAACGACACCGGATACCTACTTCTTCT	1051		
Qy	1028	CCTCAAAGCACCCAGCCAGTCAAGTGGTGACCATCTAGAAAAAGGTTTATAAAGC	1087		
Db	1052	CTTCAAAGCATCCCACTCAATCAGCTTGGTTTACCATCGTAGGAAGGATTTATAATG	1111		
Qy	1088	CTACCACTCGCAAGAGAGTATGAATGTAGCCCGTACGAAAGTTCTGCTTCTCAGTCA	1147		
Db	1112	CTACCAATTTCAAGTGAAGATTTATGAATTTGACCAATATGAAGAGTTTGTGTTCTGTCA	1171		
Qy	1148	GGTTTAAAGCGTACCCACGAATCCGATCGATCGATGCTTCTCAAGCCTCATTTCTT	1207		
Db	1172	GGTTTAAAGCGTACCCACGAATCCGATCGATGCTTCTCAAGCCTCATTTCTTCTT	1231		
Qy	1208	GTGAACAGAGAGGCTGGAGGATGGGTACACATATCTTAAATTTTGGCATATAAGAA	1267		
Db	1232	GTGAGCAAAAGGGCTTGTATTAACGATACAGCATATCCAGTTTGTCAATATAAGCAC	1291		
Qy	1268	AGCCAGAGAGTACATATTTCTATGCAAGAAATGATGACGCCAGTTTCAACAAATGTTCA	1327		
Db	1292	AGCCAGAGAGTATATATTTCCATGCAAGAAATGATGATGCCCAATTTACCAAAATGTTCA	1351		
Qy	1328	CGCTGATATAAGAAAGAAACCTCAAGTGTCTAGCAATGCTCAGCCAGCAGGCGTCT	1387		
Db	1352	CGCTGAAATATAAGAAAGAAACCTCAAGTGTCTCGAGAAAGCATCGCAAGTCAAGCGTCT	1411		
Qy	1388	GTTCTCTGATGGTACCCGCTACCTCTTGGACCTGGAAGTGTTCGACAAATCTC	1447		
Db	1412	GTTCTCGATGATATATATTTCCATGCAAGAAATGATGATGCCCAATTTACCAAAATGTTCA	1471		
Qy	1448	CCAAATTCACCGAGGAAATCCAGAGGAGTTTGAATATAAGGCTTAAACAGAAAGTGT	1507		
Db	1472	CCAACTGCACAGAGAGATCAAGAGGAGTCTGGAATATAAGAGCTTAAACAGAAAGTGT	1531		
Qy	1508	TTGGCAGTGGTGTGAGCAGTACTCTAAATATGATGAGTGGCGCGGGAAGGCTTCTGG	1567		
Db	1532	TTGGCAGTGGTGTGAGCAGTACTCTAAATATGATGAGTGGCGCGGGAAGGCTTCTGG	1591		
Qy	1568	TCAAATCTGTGGTACAAATCTTATGGCAGCTCTCGGAAACCATCTTTTAACTCAC	1627		
Db	1592	TCAAGTGTGTGATACAAATCTTCTGGCAGTCTTGTGAGAGCATCTTTTAACTCTC	1651		
Qy	1628	CAGGCCCTTCTCTTCTCATCCAGACAACTCTCTTCTATGCGACCATTTGGGCTCTGTC	1687		
Db	1652	CAGGCCCTTCTCTTCTCATCCAGACAACTCTCTTCTATGCAACAAATTTGGTGTGTC	1711		
Qy	1688	TCCCTTCTATGTTGTCTCATTTGTGTTGATCTGCGCAAAATACAAAGCAATTTAGGT	1747		
Db	1712	TCTCTTCTATTTGCTGTTTAAACCTCTGCTAATTTGTCAAGATACAAAGCAATTTAGGT	1771		
Qy	1748	ACGAGAGTCACTGTCAGATGATCCAGTGAATGATGATGATGATGATGATGATGATGATGAT	1807		
Db	1772	ATGAAGCCAGCTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1831		
Qy	1808	TTGACTTCAAGGACTATGAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1867		
Db	1832	TTGATTTCAAGAAATATGAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1891		
Qy	1868	TTGGGAAGTCTTGGGCTTGGGCTTGGGAGGTGATGAACGCGACCGCTTATGCGCA	1927		
Db	1892	TTGGGAAGTCTTGGGCTTGGGCTTGGGAGGTGATGAACGCGACCGCTTATGCGCA	1951		
Qy	1928	TTAGTAAAAAGGAGTCTCAATTTCAAGTGGCGGTGAGTGTCTTAAAGGAGAGCTGACA	1987		
Db	1952	TTAGTAAAAAGGAGTCTCAATTTCAAGTGGCGGTGAGTGTCTTAAAGGAGAGCTGACA	2011		
Qy	1988	GCTGTAAAAAGAGTCTCTATGTCGAGCTCAAAATGATGACCCACCTGGGACACCATG	2047		
Db	2012	GCTGTAAAAAGAGTCTCTATGTCGAGCTCAATTTCAAGTGGCGGTGAGTGTCTTAAAGGAGAGCTGACA	2071		
Qy	2048	ACACATCTGATCTGCTGGGGGATGACACTGTGTCAGGCGCATGCTGCTGATTTTGT	2107		

2072	AGAAATATTGTGAACCTCTCTGGGGCGTGCGACACTGTGACGACCAATTTACTTGATTTTGTG	2133
2108	AAATATTGTTGCTATGTGTGACCTCTCTCAACTACCTAAAGAAGTAAAGAAGAGAAAGTTTTCACA	2167
2132	AAATACGTGTCTATGTGTGATCTTCTCAACTATCTTAAGAAGTAAAGAAGAGAAAAATTTTCACA	2191
2168	GGACATGGACAGAGATTTTAAAGGAACATAAATTTTCAGTTCCTTACCTCTACCTTTCCAGGCAC	2227
2192	GGACTTTGGACAGAGATTTTCAAGGAAACACAAATTTTCAGTTTTTATACCCCACTTTTCCAAATCAC	2251
2228	ATTCAAAATTCACGATCCTCTGTGTTTTCAGAGAAAGTTTCAGTTTACACCCGCCCTTTGGATCAGC	2287
2252	ATCCAAATTTCAAGCATCTGGTTTCAAGAGAAAGTTTCAGATACACCCGGACTCGGATCAAA	2311
2288	TCTCAGGGTTTCAATTGGGAAATTCATTTCTCTGAAGATGAGATTGAAATATGAAGAACCCAGA	2347
2312	TCTCAGGGCTTCATGGGAAATTCATTTTCACTCTGAAGATGAAATTTGAATAATGAAGAACCCAAA	2371
2348	AGAGCTTGGCAGAAAGAAAGAGAGGAGAAAGATTTTGAACGTGTCTGACGTTTGAAGACCTCCTTT	2407
2372	AAAGGCTG-----GAAGAAAGAGGAGCAATTGAAATGTGTCTTACATTTTGAAGATCTTCTTT	2425
2408	GCTTTTGCGTACCAAGTGGCCCAAGAGGCAATGGAATTCCTCTGGAGTTTCAAGTCTGTGTCTCCACA	2467
2426	GCTTTTGCATATCAAGTTTGCCAAAGAAATGGAATTTCTTGGAAATTTAAGTCTGTGTGTTTCA	2485
2468	GAGACCTGGCAGCCACGGAAATGTGTGTGTCTACCCACGGGAAAGTGTGTGAAGATCTGTGACT	2527
2486	GAGACCTGGCCGCACAGAAAGCTGTCTGTCAACCCACGGGAAAGTGTGTGAAGATATGTGACT	2545
2528	TTGGAATCTGTGTCCGAGACATCTCTGAGCGACTTCAGACTACGTCTCAGGGGCAACGCAACGGC	2587
2546	TTGGATTTGGCTCGAGATATCATGAGTGTATTCCTCAACTATGTGTCTCAGGGGCAATCCCGCTC	2605
2588	TGCCGTGAAGTGAGATGGCAACCCAGAGACTTATTTTGAAGGATCTACACAAATCAAGAGTG	2647
2606	TGCCTGTAAAAATGGATGGCCCGCCGANAAGCCCTGTTTTGAAGGCACTACACCAATTAAGAGTG	2665
2648	ACGTCCTGCTCTACGGCATCTCTCTCTGGGAGATATTTTCACTCGGTGTGAAACCTTTACC	2707
2666	ATGTCCTGGTCATATGGAATATTTACTGTGTGGGAAATCTTCTCACTTGGTGTGNAATCCTTACC	2725
2708	CTGGCATTCCTGTCTCGACGCTAACTTCTATAAACTGATTCAGATGGATTTTAAATATGAGC	2767
2726	CTGGCATTCGGGTTTGATGCTAACTTCTACAAACTGATTCAAAAATGGATTTAAAAATGGATC	2785
2768	AGCATCTTATGCCACAGAGGGATATATCTTTGTAAATGCAATCTCTGCTGGCTTTTGGACT	2827
2786	AGCATTTTATGTCTACAGAGAAATATACATTTAATATGCAATCTCTGCTGGGCTTTTGGACT	2845
2828	CAAGGAAGCGGCCATCTCTTCCCAACCTGACTTCAATTTTATAGGATGTGAGCTGGCAGAGG	2887
2846	CAAGGAACGGCCATCTCTTCCCTAAATTTGACTTCTGTTTTTATAGGATGTGAGCTGGCAGATG	2905
2888	CAGAAGAAG-----CATGTATACAGAAATATCCATTCATTCCTTACC	2924
2906	CAGAAGAAGCGATGTATCAGAAATGTGGATGGCCGTGTTTCGGAATGCTCTCACACTTACC	2965
2925	AAAAACAGCGGCCCTCTAGCAGAGAGAGCGGGCTCAGAGC--CCAGTCTGGCACAGGCGCCAGG	2983
2966	AAAAACAGCGCGACCTTTTCAGCAGAGAGATGGAAATTTTGGGGCTACTCTCTCGCAGGCTCAGG	3025
2984	T-GAAGATTACAGAGAAAGATTTAGCGAGGAGCGCTTGGACCCCGCC-----ACCCTAG	3038
3026	TCGAAGATTTCTGTAGAGAAATTTTAGTTTAAAGCAATTCATCTCCCTCACCTATCTCCCTAA	3085
3039	CAGGCTGTAGACCGCAGAGCCCAAGATTAGCCTTCGCTCT--GAGGAAGCGCCCTACAGCG	3096
3086	CAGGCTGTAGATTACCAAAAACAAGTAATTTTCATCACTAAAAAGAAATCTATTATCAAC	3145
3097	CGTTGCTTCTGTGACTTTTCTCTAGATGTCTGTCTGCCATTATCTC-----CAAAGTGA	3149
3146	TGCTGCTTACACAGACTTTTCTCTAGAGCGGCTCTGCGTTTACTCTGTGTTTTCAAAGGGA	3205

	Qy	3150	CTTCTATAAAATCAAACCTCTCCTCGCAGGCGGGAGGCCAATATAGACTTTGTGG	3209
	Db	3206	CTTTTGTA AAAATCAAAATCATCTGTGCACAAGGAGGAGGAGCTGATAAATGAACCTTTATTG	3265
	Qy	3210	TGAGCCGCGCTCACCTCGGGGGCCTTTTCACGAGCTTGAGGGGAAAAGCCATGTATCTCGAAA	3289
	Db	3266	GAGCAITGATCTGATCTCAAGGCGCTTCTCAGGCGGGCTTGAGTGAATTTGTGTACCTGAAG	3325
	Qy	3270	TATAGTATATCTTGTTAAATACGTGAAACAAA CCAAAACCCTTTTTTTCGTAAAGGAAAGC	3329
	Db	3326	TACAGTATATCTTGTTAAATACATAAAAA CAAA-----AGCAATTTTTCGTAAAGGAGAAGC	3378
	Qy	3330	TAAATATGATTTTAAATAATCTATGTTTAAATATCTATGTAACCTTTTTCATCTATTATAG	3389
	Db	3379	TAAATATGATTTTTT---AAGTCTATGTTTAAATAATAATGTAAATTTTTTCAGCTATTATTAG	3436
	Qy	3390	TGATATATTTTATGATGCAAAATAAACCTTCTACTGTAAAAA AAAAAAAAAAAAAAAAAAAAAA	3449
	Db	3437	TGATATATTTTATGGGTGGGNAATAAATTTCTACTACAGTAAAAA AAAAAAAAAAAAAAAAAAAA	3496
	Qy	3450	AAAA 3453	
	Db	3497	AAAA 3500	
		RESULT 12		
		125170		
		LOCUS	125170	3501 bp DNA linear PAT 07-OCT-1996
		DEFINITION	Sequence 3 from patent US 5548065.	
		ACCESSION	125170	
		VERSION	125170.1 GI:1605040	
		KEYWORDS	Unknown.	
		SOURCE	Unknown.	
		ORGANISM	Unclassified.	
		REFERENCE	1 (bases 1 to 3501)	
		AUTHORS	Lemischka, I.R.	
		TITLE	Tyrosine kinase receptor human flk-2-specific antibodies	
		JOURNAL	Patent: US 5548065-A 3 20-AUG-1996;	
		FEATURES	Location/Qualifiers 1..3501 /organism="unknown"	
		BASE COUNT	1068 a 709 c 784 g 940 t	
		ORIGIN		
		Query Match	64.5%; Score 2227.2; DB 6; Length 3501;	
		Best Local Similarity	79.6%; Pred. No. 0;	
		Matches 2775; Conservative	0; Mismatches 653; Indels 56; Gaps 10;	
	Qy	8	GGCTACCGCGCTCCGGAGGCCATGCGGCGTTGGCGCAGCGCAGCGCGCGGCTGC	67
	Db	35	GGGACCCCCGGGCTCCGGAGGCCATGCCGGCTTGGC---GCGCGACGCGGGCACCGTGC	91
	Qy	68	TGCTGCTTGTGTTTGTTCAGTAAATGATCTTTCAGACCGTTTACAAACCAAGACCTGCCTG	127
	Db	92	CGCTGCTCGTGTGTTTTTCTGCAATGATATTGGGACTATTACAAATCAAGATCTGCCTG	151
	Qy	128	TGATCAAGTGTGTTTTTAAATCAGTCATCAGAAACAAATGCTCATCAGCGGGAAGCCATCAT	187
	Db	152	TGATCAAGTGTGTTTTTAAATCAATCATAAAGAACAAATGATTCATCAGTGGGGAAGTCATCAT	211
	Qy	188	CGTACCGNATGGTSCGAGGATCCCCAGAACCTCCAGTGTACCCCAGGCGCCAGAGTG	247
	Db	212	CATATCCCATGGTATACAGATCCCCCGGAACCTCCGGGTGTGGCTTGGAGACCCAGAGCT	271
	Qy	248	AAGGGA CGGTATATGAACCGGCCACCGGTGAGGTGGCCGAGTCTGGGTCCTTCAACCCCTGC	307
	Db	272	CAGGGA CAGTGTACGAAGCTGCCCTGTGGNAATGGATGATATCTGCTTCATCACACTGC	331
	Qy	308	AAGTCAAGCTGCCCA CCCCAGGGGA CCTTTCCTGCCCTCTGGGTCTTTAAGCACAGTCCC	367
	Db	332	AAGTGTGGTTCATGTC CCGAGGAACATTTCTGTCTCGGTCTTTAAGCACAGTCCCC	391

QY 368 TGGGCTGCGCAGCGGCACATTTGATTTACAAAACAGAGGAATCGTTTCCATGCGCCATCTTGA 427  
Db 392 TGAATTGCGCAGCCACATTTTGTATTTACAAAACAGAGGAGTTGTTTCCATGGTCAATTTGA 451  
QY 428 ACGTGACAGAGACCCAGGAGGAGAAATACCTACTCCATATTCAGAGCGAAACGCGCAACT 487  
Db 452 AAATGACAGAAACCCCAAGCTGAGAAATACCTACTTTTATTCAGAGTGAAGTACCAAT 511  
QY 488 ACACAGTACTGTTACAGTGAATGTAAGAGATACACAGCTGTATGTCTAAGGAGACCTT 547  
Db 512 ACACAAATATGTTTACAGTGAATGTAAGAAATACCTCTGCTTTACACATTAAGAGACCTT 571  
QY 548 ACTTTAGGAAGATGAAAACACAGGATGCACTCTCTGCATCTCCGAGGGTGTTCGGGAGC 607  
Db 572 ACTTTAGAAAATGGAANAACAGAGCGCCCTGGTCTGCAATCTGAGAGCGTTCCAGAGC 631  
QY 608 CCACTGTGGAGTGGGTGCTCTGAGCTCCACAGGGAAGCTGTAAAGAAAGAGCCCTG 667  
Db 632 CGATCGTGGAAATGGGTGCTTTGCGATTCAAGGGGGAAGCTGTAAAGAAAGTCCAG 691  
QY 668 CTGTGTTCAGAAAGAGGAAAGGTACTTCAATGATTTGTCGGAACAGACATCATAGTCT 727  
Db 692 CTGTGTGTTAAAAGAGGAAAGGTCTTCAATGAAATTTTGGGACGGACATAAGTGTCT 751  
QY 728 GTGCTAGAAATGCACTGGGCGCGGAATGCAACAGCTGTTCCACATAGATCTAAACAGG 787  
Db 752 GTGCCAGAAATGAATCTGGGAGGGAATGCAACAGCTGTTCCACATAGATCTAAATCAAA 811  
QY 788 CTCCTCAGAGCACACTGCCCGAGTTTATCTCTGAAAGTGGGGGAACCCCTTGTGGATCAGT 847  
Db 812 CTCCTCAGACCACATTTGCCAAATTTTCTTAAAGTAGGGGAACCCCTTATGGATAGGT 871  
QY 848 GTAAGGCCATCCATGTGAACATGGATTCGGGCTCACCTGGAGCTGGAAGCAAAAGCCC 907  
Db 872 GCAAAGCTGTTCATGTGAACCATGGATTCGGGCTCACCTGGGAATTAGAAAACAAAGCAC 931  
QY 908 TGGAGGAGGCGAGCTACTTTGAGATGAGTACTACTCCAAAACAGGACCAATGATTCGGA 967  
Db 932 TCGAGGAGGCGAATCTTTGAGATGAGTACTTATCAACAAACAGAACTATGATAGCGA 991  
QY 968 TTCTCTTGGGCTTTGTCTCTTCCGTGGGAAGGAACGACACCGGATATTTACACCTGTCTTT 1027  
Db 992 TTCTGTTTGTCTTTGTATCATCAGTGCAGGAAGAACGACACCGGATATCTACATTTGTCCT 1051  
QY 1028 CCTCAAAAGCACCCAGCAGCTAGCGTTGGTGACCATCTAGAAAAAGGGTTTATAAACG 1087  
Db 1052 CTTCAAAAGCATCCCAAGTCAATCAGCTTTGGTTTACCATCGTAGGAAAGGGATTTATAAATG 1111  
QY 1088 CTACCACTCGCAAGAGAGTATGAATTCACCGGTACGAAAGTTCTGCTTCTCAGTCA 1147  
Db 1112 CTACCAATTCAGTGAAGATTTAGAAATGACCAATATGAAGAGTTTGTGTTTCTGTGCA 1171  
QY 1148 GGTTTAAAGCGTACCCACGAATCCGATGCACTGGGATCTTCTCTCAAGCCTCAATTTCTTT 1207  
Db 1172 GGTTTAAAGCCTACCCACAAATCAGATGATAGTGGACCTTCTCTGGAATATCAATTTCTTT 1231  
QY 1208 GTGAAACAGAGAGCCCTGGAGGATGGGTACAGCATATCTAAATTTTGGATCATAGAAACA 1267  
Db 1232 GTGAGCAAAAGGGTCTTGATAACGGATACAGCATATCCAAAGTTTGGCAATCAAGCAACC 1291  
QY 1268 AGCCAGGAGTAGTACATATTTATGTCAGAAATATGATGATGCCCCAATTTACCCAAATGTTCA 1327  
Db 1292 AGCCAGGAGATATATTTTCCATGCGAAATATGATGATGCCCCAATTTACCCAAATGTTCA 1351  
QY 1328 CGCTGAATATAAGAAAGAAACCTCAAGTGTACCAAAATGCTCAGCAGCAGCGCGTCT 1387  
Db 1352 CGCTGAATATAAGAAAGAAACCTCAAGTGTCCAGAGCATCGGAAGTCAAGCGCTCT 1411  
QY 1388 GTTCTCTGATGCTACCCGCTACCCCTCTTGGACCTGGGAAGAGTGTTCGGAACAAATCTC 1447  
Db 1412 GTTCTCTGATGATACCCATTTACCATCTTGGACCTGGGAAGAGTGTTCAGACAGTCTC 1471

QY 1448 CCAATTTGCAGCGAGGAAATCCAGAGGAGGTGTTGGAATAAAAAGGCTAAACAGAAAGTGT 1507  
Db 1472 CCAATTTGCAGCGAGGAGATCAAGAGGAGTCTGGAATAGAAAGGCTAAACAGAAAGTGT 1531  
QY 1508 TTGGCCAGTGGGTGTCGAGCAGTACTCTAAATATAGTGAAGCCGGGAAGAGGCTTCTGG 1567  
Db 1532 TTGGACAGTGGGTGTCGAGCAGTACTCTAAACATGAGTGAAGCCATTAAGAGGTTCTCTGG 1591  
QY 1568 TCAAAATGCTGTGCGTACAAATTTCTATGGGCACTCTTTGGGAAACCATCTTTTAAACTCTAC 1627  
Db 1592 TCAAGTGTGTGTCATACAAATTTCCCTTTGGCACATCTTTGAGAGCATCTTTTAACTCTC 1651  
QY 1628 CAGGCCCCCTTCCCTTTTCATCCAAAGCAACATCTCTTATGCGACCATTTGGGCTCTGTC 1687  
Db 1652 CAGGCCCCCTTCCCTTTTCAACCAAGCAACATCTCTATCTATGCAACAAATTTGGTGTGTC 1711  
QY 1688 TCCCTCTCATTTGTTGTTCTCATTTGTTGATCTGCCACAAATACAAAAGCAATTTAGGT 1747  
Db 1712 TCCTCTTCATTTGTCGTTTAAACCTCTGTCTTAAATTTGTCAAGTACAAAAGCAATTTAGGT 1771  
QY 1748 ACAGAGTACGCTGCAGATGATCCAGGTGACTTGGCCCCCTCGATAAACGAGTACTTCTACG 1807  
Db 1772 ATGAAGCCAGCTACAGATGATGTCAGGTGACCGGCTCTCAGATATAGTACTTCTACG 1831  
QY 1808 TTGACTTCAGGGAATATGAATATGACCTTAAAGTGGGAGTTCCCGAGAGAGAACTTAGAGT 1867  
Db 1832 TTGATTTTCAGAGAAATATGAATATGATCTCAATGGGAGTTTCCAAAGAGAAATTTAGAGT 1891  
QY 1868 TTGGGAAGGCTCTGGGCTCTGGGCTTTTCGGAGGGTATGAAACGCCACGGCTTATGGCA 1927  
Db 1892 TTGGGAAGGTACTAGGATCAGGTGCTTTTGGAAAAGTATGAAACGCCAACAGCTTATGGAA 1951  
QY 1928 TTAGTAAAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTTAAAGAGAAAGCTGACA 1987  
Db 1952 TTAGCAAAAACAGAGTCTCAATCCAGGTGCGGTCAAAATGCTGAAAAGAAAAGCAGACA 2011  
QY 1988 GCTGTGAAAAGAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTCGGGAACACCATG 2047  
Db 2012 GCTCTGAAAGAGAGGCACTCATGTCAAGACTCAAGATGATGACCCAGCTGGGAAGCCAG 2071  
QY 2048 ACACATCTGTTGAATCTGCTGGGCGATGACACTGTCTCAGGCGCAGTGTACTTTGTTG 2107  
Db 2072 AGAATATTTGGAACCTGCTGGGCGGTGTCACATGTCAAGGACCAATTTACTTTGATTTTG 2131  
QY 2108 AATATTTGTTCTATGTTGAGCTCTCTCAACTACCTTAAAGTAAAGAGAGAGTTCACA 2167  
Db 2132 AATACTGTTGCTATGTTGATCTTCTCACTATCTAAGAGTAAAGAGAGAAATTTTACA 2191  
QY 2168 GGACATGGACAGAGATTTTAAAGAAACATAATTTTCACTTCTTACCTTCTTCCAGGCAC 2227  
Db 2192 GGACTTGGACAGAGATTTTCAAGGAAACAAATTTTCACTTCTTACCTTCTTCCAAATCAC 2251  
QY 2228 ATTCAAAATCCAGCATGCTGTTTCAGGAGAGTTCAGTTACACCCGCCCTTTGGATCAGC 2287  
Db 2252 ATCAAAATTCAGCATGCTGTTTCAAGAGAGTTCAGATACACCCCGGACTCGGATCAAA 2311  
QY 2288 TCTCAGGGTTCATAGTGGAAATTCATTTCTCAAGATGAGATTGAATATGAAGAACCA 2347  
Db 2312 TCTCAGGGCTTCATGGGAATTCATTTCACTCTGAGATGAAATTTGAATATGAAGAACCA 2371  
QY 2348 AGAGGCTGGCAGAGAGAGGAGGAAGATTGAACGCTGCTGAGCTTTGAAGACTCTCTTT 2407  
Db 2372 AAAGGCTG-----GAAAGAGAGGAGGACTTGAATGTGCTTACATTTGAAGATCTCTTT 2425  
QY 2408 GCTTTGCTACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAAGTCTGTGTCCACA 2467  
Db 2426 GCTTTGCTATCAAGTTGCCAAAGGAAATGGAATTTCTGGAATTTAAAGTCTGTGTTCACA 2485  
QY 2468 GAGACTGCGCAGCAAGAAATGTTGGTCAACCCAGGAGGTTGGTGAAGATCTGTGACT 2527  
Db 2486 GAGACTGCGCCCGCAGAAACGCTGTGTCAACCCAGGAAAGTGTGGAAGATATGTGACT 2545  
QY 2528 TTGGACTGCGCCCGCAGACATCTCTGAGCGACTCCAGCTTACGTCAGGGGCAACGCGCGC 2587





Db 812 CTCTCAGACACATTGCCACAATTATTCTTAAAGTAGGGAAACCTTATGGTAAGGT 871  
 Qy 848 GTAAGGCCATCCATGTGAACCATGATTCGGGCTCCTACCTGGAGCTGGAAAGCAAGCC 907  
 Db 872 GCAAGGCTGTTCATGTGAACCATGATTCGGGCTCCTACCTGGAAATAGAAAACAAAGCAC 931  
 Qy 908 TGAGGAGGGCAGCTACTTTGAGATGATGATCTACTCCACAAACAGACCATGATTCGGA 967  
 Db 932 TCAGAGGGGCACTACTTTGAGATGATGATCTACTCCACAAACAGACCATGATGATGGA 991  
 Qy 968 TTCTCTTGGCTTTGTCTCTTCCTGGGAAAGGAACGACACCGGATATTACACCTGCTCTT 1027  
 Db 992 TTCTGTCTTGTCTTTGTATCATCTAGTGGCAAGAAACGACACCGGATATTACACCTGCTCTT 1051  
 Qy 1028 CCTCAAGACACCCAGCCAGCTGAGCTTGTGATGATCTCTAGAACAAAGGGTTTATAAGC 1087  
 Db 1052 CTTCAAGCATCCAGCTCAATCAGCTTTGGTTACCATCGTAGGAAGGGATTTATAAATG 1111  
 Qy 1088 CTACCAAGCTCGCAAGAGATGATGAATTTGACCCGACGAAAGTTCTGCTTCTCAGTCA 1147  
 Db 1112 CTACCAATTCAGTGAAGATTTAATTTGACCAATATGAAGAGTTTGTCTTCTGTCTCA 1171  
 Qy 1148 GGTTTAAGCGTACCCACGAATCCGATGACGCTGGATCTTCTCAAGCTCAATTTCTT 1207  
 Db 1172 GGTTTAAGCGCTACCCACGAATCCGATGACGCTGGATCTTCTCAAGCTCAATTTCTT 1231  
 Qy 1208 GTGAACAGAGGGCTGGAGGATGGTACAGATATCTAAATTTTGGCATATAAGAA 1267  
 Db 1232 GTGAGCAAAAGGGCTTGATTAACCGGATACAGCATATCCAAATTTTGCATCATAGCACC 1291  
 Qy 1268 AGCCAGGAGATGATATTTCTATGCAAGAAATGATGACGCCAGTTCACCAAAATGTTCA 1327  
 Db 1292 AGCCAGGAGATATATTTCTATGCAAGAAATGATGACGCCAGTTCACCAAAATGTTCA 1351  
 Qy 1328 CGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAATGCTGACGAGGAGTTCGCAAGTCCCT 1387  
 Db 1352 CGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAATGCTGACGAGGAGTTCGCAAGTCCCT 1411  
 Qy 1388 GTTCTCTGATGGCTACCCGCTACCTCTTGACCTGGAAGAGTGTTCGCAAGTCCCT 1447  
 Db 1412 GTTCTCTGATGGATACCCATTAACCATCTTGACCTGGAAGAGTGTTCGCAAGTCCCT 1471  
 Qy 1448 CCAATTCGACGAGGAAATCCCAAGAGAGTTCGCAAGTGTTCGCAAGTCCCT 1507  
 Db 1472 CCAATTCGACGAGGAGATCACAGAGAGTTCGCAAGTGTTCGCAAGTCCCT 1531  
 Qy 1508 TTGGCCAGTGGGTTCGAGCAGTACTCTAAATATGATGAGTGGCGGAGGAGGCTTCG 1567  
 Db 1532 TTGGCAGTGGGTTCGAGCAGTACTCTAAATATGATGAGTGGCGGAGGAGGCTTCG 1591  
 Qy 1568 TCAATGCTGTGGTACAAATCTATGAGGAGTTCGCAAGTGTTCGCAAGTCCCT 1627  
 Db 1592 TCAATGCTGTGGTACAAATCTATGAGGAGTTCGCAAGTGTTCGCAAGTCCCT 1651  
 Qy 1628 CAGGCGCTTCCCTTTCATCAAGAGAGTTCGCAAGTGTTCGCAAGTCCCT 1687  
 Db 1652 CAGGCGCTTCCCTTTCATCAAGAGAGTTCGCAAGTGTTCGCAAGTCCCT 1711  
 Qy 1688 TCCCTTCAATGTTGTCTCAATGTTGTGATCTGCAAGTGTTCGCAAGTCCCT 1747  
 Db 1712 TCCCTTCAATGTTGTCTCAATGTTGTGATCTGCAAGTGTTCGCAAGTCCCT 1771  
 Qy 1748 ACAGAGTACGTCGAGATGATCCAGGTGCTGGCGCTTCGCAAGTGTTCGCAAGTCCCT 1807  
 Db 1772 ATGAAGCCAGCTACAGATGGTACAGTACCGGCTCCTCAGATATGATGATCTCTACG 1831  
 Qy 1808 TTGACTTCAGGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1867  
 Db 1832 TTGATTTCAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1891  
 Qy 1868 TTGGAAGGTCTCGGGCTCTGGCGCTTTCGCGAGGATGATGATGATGATGATGATGATGAT 1927  
 Db 1892 TTGGAAGGTCTAGGATCAGGTCTTTTGGAAAGTGTGATGATGATGATGATGATGATGAT 1951

Qy 1928 TTAGTAAACGGGAGTCTCAATTCAGTGGGCTGGAAGATGCTTAAAGAGAAAGCTGACA 1987  
 Db 1952 TTAGCAAAACAGAGTCTCAATTCAGTGGGCTGGAAGATGCTTAAAGAGAAAGCTGACA 2011  
 Qy 1988 GCTGTGAAAAGAAAGCTCTCATGTCGAGCTCAAAATGATGACCACTGGGACCAATG 2047  
 Db 2012 GCTGTGAAAAGAGGAGCTCATGTCAGAACTCAAGATGATGACCACTGGGAGGACG 2071  
 Qy 2048 ACAAATCTGTAATCTGCTGGGGCATGCACTGTCAGGGCCAGTGTACTTGTGATTTTG 2107  
 Db 2072 AGAATATTGTGAACCTGCTGGGGCGTGCACACTGTGAGGACCAATTTACTTGTGATTTG 2131  
 Qy 2108 AATATTGTTGCTATGTCCTCTCAACTACCTTAAGAGTAAAGAGAGAGGTTTCA 2167  
 Db 2132 AATATTGTTGCTATGTCCTCTCAACTATTAAGAGTAAAGAGAGAGGTTTCA 2191  
 Qy 2168 GGAATGGAAGAGAGTATTTTAAAGAACTAATTTTCACTTCTTACCTTCTTCCAGGCAC 2227  
 Db 2192 GGAATGGAAGAGAGTATTTTCAAGAACTAATTTTCACTTCTTACCTTCTTCCAGGCAC 2251  
 Qy 2228 ATTCAATTTCCAGCATGCTGCTGTTCAAGAAAGTTCAGTTACACCCGCTTCTGATCAGC 2287  
 Db 2252 ATCCAAATTTCCAGCATGCTGCTGTTCAAGAAAGTTCAGTTACACCCGCTTCTGATCAGC 2311  
 Qy 2288 TCTCAGGGTTCATCGGAATTTCAATTTCTTCAAGAGTGAATTTGAATATGAATAACAGA 2347  
 Db 2312 TCTCAGGGTTCATCGGAATTTCAATTTCTTCAAGAGTGAATTTGAATATGAATAACAGA 2371  
 Qy 2348 AGAGCTGGCAAGAGAGGAGGAAATTTTCAAGCTGCTGAGTTCAGTTACACCCGCTTCTT 2407  
 Db 2372 AAAGCTG-----GAAGAGAGGAGGACTTGAATGCTTCACTTGAAGATCTTCTT 2425  
 Qy 2408 GCTTTGCTACAAAGTGGCCAAAGCATGGAATTTCTGAGTTCAGTTACACCCGCTTCTT 2467  
 Db 2426 GCTTTGCTACAAAGTGGCCAAAGCATGGAATTTCTGAGTTCAGTTACACCCGCTTCTT 2485  
 Qy 2468 GAGACCTGGGAGGAGGAGTGTGTGTCACCCAGGAGGAGTGTGAGTCTGTGACT 2527  
 Db 2486 GAGACCTGGGAGGAGGAGTGTGTGTCACCCAGGAGGAGTGTGAGTCTGTGACT 2545  
 Qy 2528 TTGACTGGGAGGAGGAGTGTGTGTCACCCAGGAGGAGTGTGAGTCTGTGACT 2587  
 Db 2546 TTGATTTGGCTCGAGATATCATGAGTGTGTCACCCAGGAGGAGTGTGAGTCTGTGACT 2605  
 Qy 2588 TGCCGCTGAAAGTGGATGGGCAACCCAGAGGCTTATTTGAAGGATCTACCAATCAAGAGT 2647  
 Db 2606 TGCCGCTGAAAGTGGATGGGCAACCCAGAGGCTTATTTGAAGGATCTACCAATCAAGAGT 2665  
 Qy 2648 ACCTGCTGCTACCGCATCTTCTCTGGGAGATATTTTCACTGGGCTGTGAACCTTACC 2707  
 Db 2666 ATGCTGGTCAATGGAATATTTCTGGGAAATCTTCTCACTGGTGTGAATCTTACC 2725  
 Qy 2708 CTGGCATTTCTGCTGAGCTTAACTTCTATATAACTGATTTCAGAGTGGATTTAAATGAGC 2767  
 Db 2726 CTGGCATTTCTGCTGAGCTTAACTTCTATATAACTGATTTCAGAGTGGATTTAAATGAGC 2785  
 Qy 2768 AGCCATTTCTATGCAAGAGGATATCTTTGTAATGCAATCTGCTGGGCTTTTTCAGT 2827  
 Db 2786 AGCCATTTCTATGCAAGAGGATATCTTTGTAATGCAATCTGCTGGGCTTTTTCAGT 2845  
 Qy 2828 CAAGGAGGCGGCTTCTTCCCACTGATTTTGAAGTGTGAGTGTGAGTGTGAGTGTGAGT 2887  
 Db 2846 CAAGGAGGCGGCTTCTTCCCACTGATTTTGAAGTGTGAGTGTGAGTGTGAGTGTGAGT 2905  
 Qy 2888 CAGAAGAG-----CATGATCAAGAGGATGAGGCTGTTTTCGGAATGTCTCTCACCTTACC 2924  
 Db 2906 CAGAAGAGGATGATGAGGATGAGGCTGTTTTCGGAATGTCTCTCACCTTACC 2965  
 Qy 2925 AAAACAGGCGGCTTCTTCCCACTGATTTTGAAGTGTGAGTGTGAGTGTGAGTGTGAGT 2983  
 Db 2966 AAAACAGGCGGCTTCTTCCCACTGATTTTGAAGTGTGAGTGTGAGTGTGAGTGTGAGT 3025

Qy	2984	T-GAAGATTACACAGAAAGAGTTAGCAGGAGCCCTTGACCCCGCC-----ACCCTAG	3038
Db	3026	TCGAAGATTTCGTAGAGGAACAATTTAGTTTAAAGAACTTCATCCCTCCACCTATCCCTAA	3085
Qy	3039	CAGCTGTAGACCGCAGAGCCAAAGATTAGCCTCGCCTCT--GAGGAAGCGCCCTACAGCG	3096
Db	3086	CAGCTGTAGATTACCAAAAACAAGTTAATTTTCATCACTAAAAGAAATCTATTATCAAC	3145
Qy	3097	CGTTGCTTCGCTGGACTTTTCTCTAGATGCTGTCTGCCATTACTC-----CAAAGTGA	3149
Db	3146	TGCTGCTTACCAGACTTTTCTCTAGAACGCTCTGCGTTTACTCTTGTTTTCAAAGGGA	3205
Qy	3150	CTTCTATAAATCAAACTCTCTCGCACAGCGGAGAGCCATAATGAGACTTGTGG	3209
Db	3206	CTTTTGTAAATCAAAATCATCTGTCAAGGCGAGGAGCTGATAATGAACATTATTG	3265
Qy	3210	TGAGCCCGCTACCTCGGGGCGCTTCCACGAGCTTGAGGGGAAGCCATGATCTGAAA	3269
Db	3266	GAGCAATTGATCTGCATCCAGGCCCTTCTCAGCGCGCTTGAGTGAATTTGTGTACCTGAAG	3325
Qy	3270	TATAGTATATTCTTGTAAATACGTGAACCAAAACCCGTTTTTTTGTCTAAGGGAAAGC	3329
Db	3326	TACAGTATATTCTGTAAATACATAAACAACAA-----AGCATTTTGTCTAAGGAGAAGC	3378
Qy	3330	TAAATATGATTTTAAAAATCTATGTTTAAATACTATGTAACCTTTTTCATCTATTAG	3389
Db	3379	TAATATGATTTT---AAGTCTATGTTTAAATAATATATGTAATTTTTCAGCTATTAG	3436
Qy	3390	TGATATATTATGATCGGAATAAATCTCTACCTGTAAATAAATAAATAAATAAATAA	3449
Db	3437	TGATATATTTTATGGTGGGAATAAATAATTTCTACTACAGAAATAAATAAATAAATAA	3496
Qy	3450	AAAA 3453	
Db	3497	AAAA 3500	
RESULT 14			
LOCUS	140109	3476 bp	DNA linear PAT 13-MAY-1997
DEFINITION	Sequence 1 from patent US 5618709.		
ACCESSION	140109		
VERSION	140109.1 GI:2083114		
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 3476)		
AUTHORS	Gewirtz,A.M., Small,D. and Civin,C.I.		
TITLE	Antisense oligonucleotides specific for STK-1 and method for inhibiting expression of the STK-1 protein		
JOURNAL	Patent: US 5618709-A 1 08-APR-1997,		
FEATURES	Location/Qualifiers		
source	1..3476		
BASE COUNT	1044 a	709 c	783 g 940 t
ORIGIN	/organism="unknown"		
Query Match	63.9%; Score 2204.8; DB 6; Length 3476;		
Best Local Similarity	79.5%; Pred. No. 0;		
Matches 2752; Conservative	0; Mismatches 652; Indels 56; Gaps 10;		
Qy	8	GGCTACCGCGCTCCGGAGCCATCGCGGCGTTGGCGCAGCGCAGCGACCGCGCGCTGC	67
Db	35	GGGGACCCCGGCTCCGGAGGCATCCCGCGGTGGC-----GGCGGACGGCGGACCGTGC	91
Qy	68	TGCTGCTGCTGTTTGTTCAGTAATGATTTCTTGAGACCGTTTACAAACCAAGACCTCGCTG	127
Db	92	CGCTGCTGTTGTTTCTCGAATGATTTTGGGACTATTACAAATCAAGATCTCGCTG	151
Qy	128	TGATCAAGTGTGTTTAAATCAGTATGAGAAATGGCTCATCAGCGGGAAGCCATCAT	187
Db	152	TGATCAAGTGTGTTTAAATCAATCATAGAAACAAATGATTTCATCAGTGGGGAAGTCATCAT	211

Qy	188	CGTACCGAATGGTGGAGGATCCCCAGAAAGACCTCCAGTGTACCCGAGGCCAGAGTG	247
Db	212	CATATCCCATGTTATCAGAAATCCCGGAAGACCTCGGGTGTGGTTGAGACCCAGAGCT	271
Qy	248	AAGGACGCGTATATGAAGCGGCCACCGTGGAGGTGGCCGAGTCTGGGTCCATCAACCTGC	307
Db	272	CAGGACAGTGTACGAAGCTGCCGCTGTGGAAGTGGATGTATCTGCTTCCATCAACCTGC	331
Qy	308	AAGTGCAGCTGCCACCCAGGGGACCTTTCTCGCTCTGGGTCTTTAAGACAGCTGCC	367
Db	332	AAGTGTGCTGTCATGCCCCCAGGGAACATTTCTGCTGTGGTCTTTAAGCAAGCTGCC	391
Qy	368	TGGGCTGCCAGCGCATTTTGTATTTACAAAACAGAGGAATCGTTTCCATGGCCATCTCA	427
Db	392	TGAATTTGCCACCCACATTTTGAATTTACAAAACAGAGGAGTTGTTTCCATGGTCAATTTGA	451
Qy	428	ACGTGACAGAGACAGGAGGAGAAATACCTACTCATATTTCAGAGCGGAACCGCCAACT	487
Db	452	AAATGACAGAAACCCAGCTGGAGAAATACCTACTTTTATTCAGAGTGAAGTACCAAT	511
Qy	488	ACACAGTACTGTTCAAGTGAATGTAAGAGATACACAGCTGTATGTCTAAGAGACCTT	547
Db	512	ACACAAATTTGTTTACAGTGTAGTATAAGAAATACCTGCTTTACACATTTAAGAAAGACCTT	571
Qy	548	ACTTTAGAAAGATGAAACCCAGGATGCATGCTCTGCTCTCCGAGGGTGTTCGGAGC	607
Db	572	ACTTTAGAAAAATGAAAAACCCAGGACGCCCTGCTGCTATATCTGAGAGCGTTCCAGAGC	631
Qy	608	CCACTGTGAGTGGTGTCTCTGCAGCTCCCAACAGGAAAGCTGTAAAGAAAGAGCCCTG	667
Db	632	CGATCGTGAATGGGTGCTTTGCGATTACAGGGGGGAAGCTGTAAAGAAAGAAAGTCCAG	691
Qy	668	CTGTTGTGAGAAAGAGGAAAGGTTACTTTCATGAGTGTTCGGAACAGACATCAGATGCT	727
Db	692	CTGTTGTTAAAAAGAGGAAAGGTTTCTCATGAATTTATTTGGGACGACATAAGGTGCT	751
Qy	728	GTGCTAGAAATGCACTGGGCGCGGAATGCAACCAAGCTGTTTCAACATAGATCTTAAACAGG	787
Db	752	GTGCCAGAAATGAATGGGCGAGGAAATGCACAGGCTGTTTCAATAGATCTTAAATCAAA	811
Qy	788	CTCCTCAGAGCACACTGCCCCAGTTTATTCCTGAAAGTGGGGAAACCTTTGTGATCAGGT	847
Db	812	CTCCTCAGACCACTTGCCCAATTTATTTCTTAAAGTAGGGGAACCTTTATGATAGGT	871
Qy	848	GTAGGECATCCATGTGAACCAATGGATTGCGGCTCACCTGGGAGCTGGAAGACAAAGCCC	907
Db	872	GCAAAGCTGTTTCAATGTGAACCAATGGATTGCGGCTCACCTGGGAAATTAGAAAACAAAGCAC	931
Qy	908	TGGAGAGGGCAGCTACTTTTGAAGTGAAGTACCTTCTCAACAAAGACCAATGATTCGGA	967
Db	932	TCGAGGAGGGCAACTACTTTGAGATGAGTACCTATTTCACAAACAAACAGAACTATGATACGA	991
Qy	968	TTCTCTTGGCTTTGTTGCTTCCGTTGGGAGGAACGACACCGGATATTACCTGCTCTT	1027
Db	992	TTCTGTTGCTTTTGTATCATCAGTGGCAAGAAAACGACACCGGATATTACACTGTTCTCT	1051
Qy	1028	CCTCAAAAGCACCCAGCCAGCTCAGCTGCTGCTGACCAATCTCTAGAAAAGGTTTATAAACG	1087
Db	1052	CTTCAAGCATCCAGTCAATCAGTTTGTGTACCATCGTAGAAAAGGATTTATAATG	1111
Qy	1088	CTACAGCTCGCAAGAAGATGAAATTTGACCCGCTACCGGATCGAAAGTTCTGCTCTCAGTCA	1147
Db	1112	CTACCAATTTCAAGTGAAGATTATGAAATTTGACCAATATGAAGAGTTTGTGTTTCTGTCA	1171
Qy	1148	GTTTTAAAGCTACCCACGAATCCGATGCAGTGGATCTTCTCTCAAGCTCATTTCTT	1207
Db	1172	GTTTTAAAGCTTACCCACAAATCAGATGTACGTGGACCTTCTCTCGAAAATCATTTCTT	1231
Qy	1208	GTGACAGAGAGCGCTGGAGGATGGGTACAGATATCTAAATTTTTCGATCATTAAGACA	1267
Db	1232	GTGACAAAGAGGTCTTTGATAACGGATACAGATATCTCAAGTTTTCGAATCATAGCAC	1291

QY 1268 AGCCAGGAGAGTACATATCTTATGCGAGAAAATGATGACGCCAGTTCCACCAAAATGTTCA 1327  
D 1292 AGCCAGGAGAGATATATTTCCATGCGAGAAATGATGATGCCAAATTTACCAAAATGTTCA 1351  
QY 1328 CGCTGAATATAGAAAAGAAACCTCAAGTGTAGCAAAATGCTTCAGCCAGCAGCGGCTCT 1387  
D 1352 CGCTGAATATAGAAAAGAAACCTCAAGTGTAGCAAAATGCTTCAGCCAGCAGCGGCTCT 1411  
QY 1388 GTTCTCTGATGGCTACCCGCTACCCCTTTGGACCTGGAAGAGTGTTCGGAACAATCTC 1447  
D 1412 GTTCTCGATGGATACCCATTTACCATCTTGGACCTGGAAGAGTGTTCGGAACAATCTC 1471  
QY 1448 CCAATTGACGGAGAAATCCAGAAAGAGTTTGGAAATAAAAGGCTAAACAGAAAAGTGT 1507  
D 1472 CCAATGACAGAAAGATACAGAAAGAGTGTGGAATAGAAAGGCTTAACAGAAAAGTGT 1531  
QY 1508 TTGGCCAGTGGGTGTGAGCAGTACTCTAAATATGAGTGAAGGCCGGAAAGGCTTCCTGG 1567  
D 1532 TTGGCAGTGGGTGTGAGCAGTACTCTAAATATGAGTGAAGGCCGGAAAGGCTTCCTGG 1591  
QY 1568 TCAAAATGCTGTGCTGATCAATCTTATGGGCAAGTCTTCGGAACCAATCTTTTAAACTCAC 1627  
D 1592 TCAAGTGTGTGATCAATCTTATGGCAGTACTCTTGGCAGTCTTGTGAGAGGATCTCTTAAACTCTC 1651  
QY 1628 CAGGCCCTTCCCTTTCATCCAAAGACAACTCTCTTCTATGCGACCATTTGGCTCTGTC 1687  
D 1652 CAGGCCCTTCCCTTTCATCCAAAGACAACTCTCTTCTATGCGACCATTTGGCTCTGTC 1711  
QY 1688 TCCCTTTCATTTGTTCTCATTTGTTGATGTCGCCAACAAATACAAAAGCAATTTAGGT 1747  
D 1712 TCCCTTTCATTTGTTCTTAAACCTCTTAATTTGTGCAAGTACAAAAGCAATTTAGGT 1771  
QY 1748 ACAGAGTCAAGTGTGAGATGATCCAGTGTGATGAGTGGGAGTTCGAGAGAGACTTAGGT 1807  
D 1772 ATGAAAGCAGGCTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1831  
QY 1808 TTGACTTCAAGGACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1867  
D 1832 TTGATTTCAAGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1891  
QY 1868 TTGGAGAGTCTGGGCTGCGCTTTTCGGAGGATGATGATGATGATGATGATGATGATGATGAT 1927  
D 1892 TTGGAGAGTCTAGGATCAGGCTGCTTTTGGAAAGTGTGATGATGATGATGATGATGATGATGAT 1951  
QY 1928 TTAGTAAACGGAGTCTCAATTCAGTGGCGGTGAAAGTGTGATGATGATGATGATGATGATGAT 1987  
D 1952 TTAGTAAACGGAGTCTCAATTCAGTGGCGGTGAAAGTGTGATGATGATGATGATGATGATGAT 2011  
QY 1988 GCTGTGAAAAGAGCTCTCATGTCGGAGCTCAAAATGATGATGATGATGATGATGATGATGAT 2047  
D 2012 GCTGTGAAAAGAGCTCTCATGTCGGAGCTCAAAATGATGATGATGATGATGATGATGATGAT 2071  
QY 2048 ACAACATCGTGAATCTGTGGGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 2107  
D 2072 AGAATATTTGAAACCTGCTGGGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 2131  
QY 2108 AATATTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2167  
D 2132 AATATTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2191  
QY 2168 GGACATGGACAGAGATTTTAAAGAACATATTTCACTTCAACTATCAAGAGTAAAGAGAGAAATTTCA 2191  
D 2192 GGACATGGACAGAGATTTTAAAGAACATATTTCACTTCAACTATCAAGAGTAAAGAGAGAAATTTCA 2227  
QY 2228 ATTCAAATTCAGATGCTGCTGTTTCAAGAGAGTTCAGTATACCCCGCTTGGATCAGC 2287  
D 2252 ATTCAAATTCAGATGCTGCTGTTTCAAGAGAGTTCAGTATACCCCGCTTGGATCAGC 2311  
QY 2288 TCTCAGGGTTCATGGGAATTTCAATTCATTTGATGATGATGATGATGATGATGATGATGATGAT 2347  
D 2312 TCTCAGGGTTCATGGGAATTTCAATTCATTTTCACTTCAAGTAAATGATGATGATGATGATGATGAT 2371  
QY 2348 AGAGGCTGGCAGAGAGAGAGAGAGTGTGAACTGCTGACGTTTGAAGACCTCTCTT 2407

D 2372 AAAGGCTG-----GAAGAGAGGAGGACTTGAATGCTTACATTTTGAAGATCTCTTT 2425  
QY 2408 GCTTTGGTACCAAGTGGCCAAAGCATGGAATTCCTGGAGTTCAGATTCGAATGCTGTGTCCACA 2467  
D 2426 GCTTTGATATCAAGTGTGCAAAAGGAATTTCTGGAATTTAAGTGTGTGTTCACA 2485  
QY 2468 GAGACCTGGCAGCAGCAAAATGTTGTGTTCACCAAGGAGGAGTGTGAGATCTGTGACT 2527  
D 2486 GAGACCTGGCAGCAGCAAAATGTTGTGTTCACCAAGGAGGAGTGTGAGATCTGTGACT 2545  
QY 2528 TTGGACTGGCCCGAGACATCTTGAAGCAGTCTTGAAGCAGTCTTGAAGCAGTCTTGAAGCAGTCT 2587  
D 2546 TTGGACTGGCCCGAGACATCTTGAAGCAGTCTTGAAGCAGTCTTGAAGCAGTCTTGAAGCAGTCT 2605  
QY 2588 TGCCGGTGAAGTGGATGGCAGCAGGAGTATTTGAAGGAGTCTACCAATCAAGAGTG 2647  
D 2606 TGCCGGTGAAGTGGATGGCAGCAGGAGTATTTGAAGGAGTCTACCAATCAAGAGTG 2665  
QY 2648 ACCTCTGGTCTTACGGCATCTTCTCTGGGAGATATTTTCACTGGGAGTGAACCTTACC 2707  
D 2666 ATCTCTGGTCTTACGGCATCTTCTCTGGGAGTATTTTCACTGGGAGTGAACCTTACC 2725  
QY 2708 CTGGCATCTCTGTGAGCAGTAACTTCTTATAAACTGATTCAGAGTGGATTTTAAATGAGC 2767  
D 2726 CTGGCATCTCTGTGAGTAACTTCTTATAAACTGATTCAGAGTGGATTTTAAATGAGC 2785  
QY 2768 AGCCATCTTATGCCACAGAGGATATCTTGAATGCAATCTCTCTGGGCTTTTGAAT 2827  
D 2786 AGCCATCTTATGCCACAGAGGATATCTTGAATGCAATCTCTCTGGGCTTTTGAAT 2845  
QY 2828 CAAGAGCGGCGCATCTTCTTCCCACTGATTCATTTTAGGATGTGAGTGTGAGGAGG 2887  
D 2846 CAAGAGCGGCGCATCTTCTTCCCACTGATTCATTTTAGGATGTGAGTGTGAGGAGG 2905  
QY 2888 CAGAGAGAG-----CATGTATCAGAAATCCATCATCTTACC 2924  
D 2906 CAGAGAGAGGATGATCAGATGTCAGATGTCGCGGTTTCGGAATGTCCTCAGACCTTACC 2965  
QY 2925 AAAACAGGCGGCTCTCAGCAGAGAGGCGGCTCAGAGC-CCAGTCCGCAAGCGGCGAGG 2983  
D 2966 AAAACAGGCGGCTCTCAGCAGAGAGGCGGCTCAGAGC-CCAGTCCGCAAGCGGCGAGG 3025  
QY 2984 T-GAAGATTCAGAGAGAGAGGATTTAGCAGAGAGGCTTGGACCC-----GGCCACCTAG 3038  
D 3026 TCGAAGATTCAGAGAGAGGATTTAGTGTGAGGAGCTTCTCTCTCCCTCCACCTATCCCTAA 3085  
QY 3039 CAGCTGTAGAGCGCAGAGCGCAAGATTTAGCCTCGCTCT- -GAGGAAGCGCCTCAGCG 3096  
D 3086 CAGCTGTAGAGCGCAGAGCGCAAGATTTAGTGTGAGGAGCTTCTCTCTCCCTCCACCTATCCCTAA 3145  
QY 3097 CGTTGCTTCTGAGGAGTCTCTAGATGCTGTGCTTCTTCTC-----CAAAGTGA 3149  
D 3146 TGCTGCTTCTCAGAGCTTCTCTAGAGCGGCTCTGCGTCTTCTCTGTTTCAAAGGGA 3205  
QY 3150 CTCTTATAAAATCAAACTCTCTCGCAGCGGAGAGCGCAATTAATGAGACTTGTGAG 3209  
D 3206 CTCTTATAAAATCAAACTCTCTCGCAGCGGAGAGCGCTGATTAATGAGACTTGTGAG 3265  
QY 3210 TGAGCCCGCTTACCTGCGGCGCTTCTTCCAGCAGCTTGAAGGAGAGCGCAATGATCTGAAA 3269  
D 3266 GAGCATTTGATCTGATCCAGGCGCTTCTCAGGCGGCTTGAAGTGTGTACTCTGAG 3325  
QY 3270 TATGATATTTCTTGTAAAATCTATGTTTAAATACTATGATTAATCTTTTCAATCTATTAG 3329  
D 3326 TACAGTATTTCTTGTAAAATCTATGTTTAAATACTATGATTAATCTTTTCAATCTATTAG 3378  
QY 3330 TAAATATGATTTTAAATACTATGTTTAAATACTATGATTAATCTTTTCAATCTATTAG 3389  
D 3379 TAAATATGATTTT- -AAGTCTATGTTTAAATACTATGATTAATCTTTTCAATCTATTAG 3436  
QY 3390 TGATATTTTATGAGTGAAGTAAATCTTCTACTGTAAA 3429



```
QY 1748 ACAGAGTCTAGCTCAGATGATCAGGTGACTGGCCCCCTGGATAAACGAGTACTTCTACG 1807
Db 1772 ATGAAGCCAGCTACAGATGTAAGGTGACCGGCTCCTCAGATAATGAGTACTTCTACG 1831
QY 1808 TTGACTTCAGGGAATATGAATATGACCTTAAGTGGAGTTCCTCCAGAGAGAACTTAGAGT 1867
Db 1832 TTGATTTACAGAGATATGAATATGATCTCAATGGGAGTTCCAGAGAGAAATTTAGAGT 1891
QY 1868 TTGGGAAGTCTGGGCTCTGGCGCTTTTCGGAGGGTATGAAACGCCACGGCCTATGGCA 1927
Db 1892 TTGGGAAGTCTAGGACTAGGACTAGGCTGCTTTGGAAAAGTATGAACGCAACAGCTTATGAA 1951
QY 1928 TTAGTAAAACGGAGTCTCAATTCAGGTGGGGTGAAGATCTAAAGAGAAAGCTGACA 1987
Db 1952 TTAGCAAAAACGGAGTCTCAATCAGGTGGCGTCAAAATGCTGAAAAGAAAACGACACA 2011
QY 1988 GCTGTGAAAAGAGCTCTCATGTGGAGCTCAAAATGATGACCCACTGGGACACCATG 2047
Db 2012 GCTGTGAAAGAGGCACTCATGTGAGACTCAAGATGATGACCCAGCTGGGAAGCCACG 2071
QY 2048 ACAACTCGTGAATCTCTGGGGGCATGCACTGTGAGGCCCACTGTACTTTGATTTTGG 2107
Db 2072 AGAATATTGTGAACCTGCTGGGGCGTGCACTGTGAGGACCAATTTACTTGAATTTTG 2131
QY 2108 AATATTGTTGCTATGGTGACTCTCTCAACTACCTAAGATGATAAGAGAGAGTTTCA 2167
Db 2132 AATACTGTTGCTATGGTGATCTCTCAACTATCTAAGAGTAAAGAGAGAAAAATTTCA 2191
QY 2168 GGACATGACAGAGATTTTAAAGAAACATAATTTCACTTCTTACCTTCTTCCAGGCAC 2227
Db 2192 GGAATTTGGACAGAGATTTTCAAGAAACATAATTTCACTTTTACCCCACTTTTCAATC 2251
QY 2228 ATTCAAATTCAGATGCTGGTTCAGAGAAAGTTCACTTACACCCGCCCTTGGATCAGC 2287
Db 2252 ATCCAAATTCAGATGCTGGTTCAGAGAAAGTTCACTTACACCCGCCCTTGGATCAGC 2311
QY 2288 TCTCAGGTTCAATGGGAATTCATTTCAATCTGAGATGATGATGATGATGATGATGATG 2347
Db 2312 TCTCAGGTTCAATGGGAATTCATTTCACTCTGAGATGATGATGATGATGATGATGATG 2371
QY 2348 AGAGGCTGGCAGAGAGAGAGAGATTTGAACGTGCTGAGCTTTTGAAGAGCTCTCTTT 2407
Db 2372 AAGGCTG-----GAAGAGAGAGAGACTTGAATGTGCTTACATTTGAAGATCTCTTT 2425
QY 2408 GCTTTCGTAACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAAGTCTGTGTTCACA 2467
Db 2426 GCTTTCGATATCAAGTTGCCAAAGGAATGGAATTTCTGGAATTTAAGTCTGTGTTCACA 2485
QY 2468 GAGACCTGGCAGGACAGGAATGTGTGTGTCACCAACGGGAAGGTGGTGAAGATCTGTGACT 2527
Db 2486 GAGACCTGGCCGCGCAGGAAGCTGTTGTCAACCAACGGGAAGGTGGTGAAGATGATGACT 2545
QY 2528 TTGGACTGGCCCGAGACATCTCAGCGACTCCAGCTACGTCAGGGGCAACGCAACGGC 2587
Db 2546 TTGGATTTGCTCGAGATATCAAGATGATTCACCACTATGTTGTGAGGGCAATGCCCCG 2605
QY 2588 TGCCGCTGAAGTGGATGGCAACCGAGAGCTTATTTGAAGGATCTACACAATCAAGAGTG 2647
Db 2606 TGCTGTAAAAATGGATGGCCCCGAAAGCTGTTTGAAGGCACTACACCAATTAAGAGTG 2665
QY 2648 AGCTGTGGTCTTACGGATCTCTCTGAGAGATATTTTCACTGGGTGTGAACCCCTTACC 2707
Db 2666 ATGTCTGGTCAATGGAATATTTACTGTGGGAATCTTCTCACTTGGTGTGAATCTCTTACC 2725
QY 2708 CTGGCAATCTCTGTCAGCGCTTAACTCTATAACTGATTCAGATGATTTTAAATGGAGC 2767
Db 2726 CTGGCAATCCGGTGTATGCTAACTCTCAAACTGAATCAAAATGGATTTTAAATGGATC 2785
QY 2768 AGCCATTTCTATGCCACAGAGGGATATCTTTGTAATGCAATCCTGCTGGCTTTTGACT 2827
Db 2786 AGCCATTTTATGCTACAGAGAAATATACATTAATGCAATCTGCTGGCTTTTGACT 2845
```

```
QY 2828 CAAGGAAGCGGCATCTCTCCCAACCTGACTTCATTTTAGGATGTCAGTGGCAGAGG 2887
Db 2846 CAAGGAAGCGGCATCTCTCCCAACCTGACTTCATTTTAGGATGTCAGTGGCAGATG 2905
QY 2888 CAGAAGAAG-----CATGTATCAGAACATCCATCCATCTTACC 2924
Db 2906 CAGGAAGCGATGATCAGAAATGTGATGGCCGTGTTTCGAAATGTCTTCAACCTTACC 2965
QY 2925 AAAACAGCGCGCCCTCAGCAGAGAGCGGCTCAGAGC-CCAGTCGCCACAGCGCCAGG 2983
Db 2966 AAAACAGCGCACCTTTTCAGCAGAGAGATGATTTTGGGCTACTCTCTCGCAGGCTCAGG 3025
QY 2984 T-GAAGATTACAGAGAAAGATTAGCAGAGGCGCTTGGACCC-----CGCCACCTTAG 3038
Db 3026 TCGAAGATTCTGAGAGGAACAATTTAGTTTAAAGACTTTCATCCCTCCACCTATCCCTAA 3085
QY 3039 CAGGCTGTACACCGCAGAGCGCAAGATTAGCCTCGCCTCT--GAGGAAGCGCCCTACAGCG 3096
Db 3086 CAGGCTGTAGATTACCAAAAACAAGATTAAATTTTCATCACTAAAGAAAATCTTATTCAC 3145
QY 3097 CGTTGCTTCTGGACTTTTCTTAGATGCTGTGCAATTTACTC-----CAAAGTGA 3149
Db 3146 TGCTGCTTCAACAGACTTTTCTTAGAGCGCTCTGCTTTACTCTTGTTTTCAAGGGA 3205
QY 3150 CTTCTATAAAATCAAACTCTCTCGCACAAGCGGAGAGCCAAATAATGAGACTTGTGG 3209
Db 3206 CTTTGTGTAATCAAAATCATCTCTGTCAGAGGCGAGGAGCTGATAATGAACTTTATTG 3265
QY 3210 TGAGCCGCGCTACCTGGGGCGCTTTCACAGAGCTTGAGGGGAAAGCCATGATCTGAAA 3269
Db 3266 GAGCATTTGATCTGCATCCAAAGCGCTTCTCAGGCGGCTTGAGTGAATTTGTACTCGAAG 3325
QY 3270 TATAGTATATTTCTGTAAATACGTGAACAAACCAACCCGTTTTTTTGTAAAGGGAAGC 3329
Db 3326 TACAGTATATTTCTGTAAATACATAAAACANA-----AGCATTTTGTCTAAGGGAAGC 3378
QY 3330 TAAATAGATTTTAAAAATCTATGTTTTTAAAAATCTATGTAATCTTTTCTATTTAG 3389
Db 3379 TAATATGATTTTTT--AGTCTATGTTTTTAAAAATAATGTAATTTTTCAGCTATTTAG 3436
QY 3390 TGNATATTTTATGGATGGAATAAATTTTCTACT 3424
Db 3437 TGATATTTTATGGGTGGGAATAAATTTTCTACT 3471
```

Search completed: August 28, 2003, 01:02:04  
Job time : 16115.1 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: August 27, 2003, 15:17:42 ; Search time 8334.07 Seconds  
(without alignments)  
10069.907 Million cell updates/sec

Title: US-09-919-408A-1  
Perfect score: 3453  
Sequence: 1 GCGCGCTGGTACCGCGGC.....AAAAAAAAAAAAAAAAAAAA 3453

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pin:\*

20: em\_gss\_vit:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2080.8	60.3	3647	11	BC036028 Homo sapi
2	1229.8	35.6	3761	11	AK045865 Mus muscu
3	655.8	19.0	765	9	AI323253
4	556.8	16.1	618	9	AA120050 mp88b02.i

5	531.8	15.4	604	10	BF523018
6	502.4	14.5	747	9	AI323643
7	495.4	14.3	500	13	BU696796
8	486.6	14.1	522	10	BB644407
9	452.6	13.1	621	12	BI360262
10	446.6	12.9	658	14	BY727686
11	445	12.9	445	10	BE51447
12	432.8	12.5	714	12	BI461248
13	428.4	12.4	597	13	BU695104
14	426.2	12.3	515	13	BQ556177
15	361.6	10.5	403	12	EM117950
16	354.4	10.3	507	12	BM484050
17	340.8	9.9	460	9	AV713950
18	333.6	9.7	3312	11	AK004947
19	333.6	9.7	3616	11	AK080326
20	333.6	9.7	3687	11	AK079247
21	332	9.6	3822	11	BC050024
22	327.4	9.5	3687	11	AK076215
23	324.8	9.4	3883	11	AK046795
24	324.8	9.4	5174	11	BC026713
25	314	9.1	394	10	BGI45808
26	313.8	9.1	3887	11	AK047010
27	298.8	8.7	451	14	CB787035
28	294	8.5	457	9	AA996414
29	279.4	8.1	332	14	CB605696
30	277.6	8.0	357	9	AI556371
31	273	7.9	424	9	AI180146
32	266	7.7	502	10	BF557871
33	265.8	7.7	455	12	BM147250
34	254	7.4	568	10	BE679272
35	253	7.3	318	14	CB548474
36	248.2	7.2	375	9	AA851570
37	245.8	7.1	939	13	BX371699
38	244.8	7.1	660	10	BE047875
39	235.8	6.8	724	14	BY764070
40	234.8	6.8	640	14	BY749257
41	233	6.7	449	13	BY545844
42	227.8	6.6	386	9	AI178876
43	227.6	6.6	407	9	AV596131
44	226.6	6.6	739	10	BF159205
45	226.6	6.6	779	14	CA511669

ALIGNMENTS

RESULT 1  
BC036028  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REMARK  
COMMENT

BC036028 3647 bp mRNA  
Homo sapiens, clone IMAGE:5272266, mRNA.  
BC036028  
BC036028.1 GI:23273564  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Rutheria; Primates; Catarrhini; Hominidae; Homo.  
Strausberg, R.  
Direct Submission  
Submitted (31-JUL-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki  
Toshiyuki and Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
Center, Stanford University School of Medicine, Stanford, CA 94305

BF523018 UI-R-C2p-  
AI323643 mp88b02.x  
BU696796 LL2in1112  
BB644407 BB644407  
BI360262 387099 MA  
BY727686 BY727686  
BE51447 UI-M-BH3-  
BI461248 603206574  
BU695104 LL2in1190  
BQ556177 H4038E08-  
BM117950 L0858E03-  
BM484050 537756 MA  
AV713950 AV713950  
AK004947 Mus muscu  
AK080326 Mus muscu  
AK079247 Mus muscu  
BC050024 Mus muscu  
AK076215 Mus muscu  
AK046795 Mus muscu  
BC026713 Mus muscu  
BGI45808 uu9807.Y  
AK047010 Mus muscu  
CB787035 AMGNNUC.N  
AA996414 UI-R-CO-h  
CB605696 AMGNNUC:M  
AI556371 UI-R-C2p-  
AI180146 EST223884  
BF557871 UI-R-CO-h  
BM147250 TCAAAPQ10  
BE679272 GG61C06.Y  
CB548474 AMGNNUC:M  
AA851570 EST194338  
BX371699 BX371699  
BE047875 t240d09.Y  
BY764070 BY764070  
BY749257 BX749257  
BY545844 BY545844  
AI178876 EST222558  
AV596131 AV596131  
BF159205 601766656  
CA511669 UI-R-FJ0-



Db 1790 TTTAGGTATGAAGCCAGCTACAGATGGTACAGGTGACCGGCTCCTCAGATATATGATGAC 1849  
Qy 1801 TTCTACGTTGACTTTCAGGGACT-ATCAATATGACCTTAAAGTGGAGTTCCTCCAGAGAGAA 1859  
Db 1850 TTCTACGTTGATTTACAGAGAAATATGATATGATCTCAATGGAGTTTCCAGAGAGAA 1909  
Qy 1860 CTTAGAGTTTGGAAAGTCTCGGGTCTCGGCTTTCGGAGGGTATGAAAGCCACCGC 1919  
Db 1910 TTTAGAGTTTGGAAAGTCTAGAGTCAAGTCTTTCGAAAGTGTGAACGCAACAGC 1969  
Qy 1920 CTATGCCATTACTAAACCGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTTAAAGAGAA 1979  
Db 1970 TTATGAAATTAGCAAAACAGGAGTCTCAATCAGGTTCGCGTCAAAATGCTGAAGAGAA 2029  
Qy 1980 AGCTCAGAGCTGTGAAAGAGAGTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGG 2039  
Db 2030 AGCAGACAGCTCTGAAGAGAGGCACTCATGTCAAGACTCAAGATGATGACCCAGCTGG 2089  
Qy 2040 ACACCATGACAACTCGTGAATCTGCTGGGGATGACACTGTGAGGGCCAGTGTACTT 2099  
Db 2090 AAGCCACGAGATAATGTGAACCTGCTGGGGCGTGACACTGTCAAGGACCAATTTACTT 2149  
Qy 2100 GATTTTGAATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2159  
Db 2150 GATTTTGAATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2209  
Qy 2160 GTTTCACAGGACATGACAGAGATTTTAAAGAAACATAATTTCACTTCTTACCTACTTT 2219  
Db 2210 ATTTTCACAGGACTTGACAGAGATTTTCAAGAAACATAATTTCACTTCTTACCTACTTT 2269  
Qy 2220 CCAGGCACATTCAAATTTCCAG----- 2240  
Db 2270 CCAATCAGATCCAAATTCAGTAAAGAAATGAGCTTTTACAAAGCAAACTGGAAGAAA 2329  
Qy 2241 ----- 2240  
Db 2330 GAAGGATGTTGAACGTTTACGGGACTCTCGGAGATCTGTATTATGTGAGGAGAAAGTG 2389  
Qy 2241 -----CATGCTGGTTTACAGAGAGTTCAGTT 2267  
Db 2390 GGCTGAGCTCAGAAACCAAGAAATGAGATCGATCATGCTGTTTCAAGAGAGTTTCAGAT 2449  
Qy 2268 ACACCGGCTTGGATCAGCTCTCAGGGTTCAATGGGAAATTCATTTCTGAAGATGA 2327  
Db 2450 ACACCGGACTCGGATCAATCTCAGGCTTCATGGGAAATTCATTTCACTCTGAAGATGA 2509  
Qy 2328 GATTGAATGAACCAAGAGGCTGGCAGAGAGAGAGAGAGATTTGAACGTGCT 2387  
Db 2510 AATTGAATGAACCAAAAAAGGCTG-----GAAGAGAGGAGGACTTGAATGTGCT 2563  
Qy 2388 GAGTTTGAAGACTCTTTGCTTTCGTTACCAAGTGGCCAGAGCATGGAATTCCTGGA 2447  
Db 2564 TACATTGAAGATCTTCTTGTCTTTCATATCAAGTTGCAAGAGGAAATGGAATTCCTGGA 2623  
Qy 2448 GTTCAAGTCTGTGTTCCACAGAGACTGGCAGCAGGAATGTGTTGTTCAACCCACGGAA 2507  
Db 2624 ATTTAAGTCTGTGTTTACAGAGACTTGGCGCCAGAGAGCTGTTGTCAACCCACGGAA 2583  
Qy 2508 GTGTGTGAAGATCTGTGACTTTGGACTGGCCCGAGACATCTCAGGAGCTCCAGTACGT 2567  
Db 2684 AGTGTGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGAGTGAATTCCAACTATGT 2743  
Qy 2568 CGTCAGGGCAACGACGCTGCGGTGAAGTGAATGGACCCAGAGACTTATTTGAAG 2627  
Db 2744 TGTCAAGGGCAATGCCCGCTGCTGCTTAAATGAGTGGCCCGGAAAGCTGTTTGAAG 2803  
Qy 2628 GATCTACAAATCAAGAGTGAAGTCTGTTCTAGGCATCTTCTGAGAGATATTTTC 2687  
Db 2804 CATCTACAAATTAAGAGTGAATGCTGTTGATATGAATATTTACTGTGGAAATCTTCTC 2863  
Qy 2688 ACTGGGTGAACCCCTTACCTGGCATTCCTGTGACGCTTAACTTCTATAAATGATTCA 2747  
Db 2864 ACTTGTGTGAATCCCTTACCTGGCATTCGGTTGATGCTTAACTTCTACAACTGATTCA 2923

Qy 2748 GAGTGGATTTAAATGGAGCAGCCATTCTATGCCACAGAGGATATACTTTGTTAATGCA 2807  
Db 2924 AAATGGATTTAAATGGATCAGCCATTTTATGCTACAGAGAAATATATCATATAATGCA 2983  
Qy 2808 ATCTCTGCTGGGCTTTTGACTCAAGAGAGCGGCATCTCTTCCCAACTGACTTCATTTT 2867  
Db 2984 ATCTCTGCTGGGCTTTTGACTCAAGAGAGCGGCATCTCTTCCCTAATTTGACTTCGTTT 3043  
Qy 2868 AGGATGTTCAGCTGGCAGAGGCAGAGAAAG-----CATGTATC 2904  
Db 3044 AGGATGTTCAGCTGGCAGATGTCAGAGAGCGGATGATCAGAAATGTGGATGGCGGTGTTT 3103  
Qy 2905 AGAAATCATCATTCATCTACCAAAACAGGCGGCCCTCAGCAGAGAGCGGGCTCAGAGC- 2963  
Db 3104 GGAATGTCTCACACCTTACCAAAACAGGCGACCTTTTTCAGCAGAGAGATGGAATTTGGGCT 3163  
Qy 2964 CAGTGGCCACAGCCCGAGT-AGAGATTCACAGAGAGAGAGATTTAGCGAGGAGGCTTG 3022  
Db 3164 ACTCTCTCGCAGGCTCAGGTGGAAGATTCGTAGAGGAAATTTAGTTTAAAGGACTTC 3223  
Qy 3023 GACCC-----CGCCACCTAGCAGGCTGTAGACCGCAGAGCAAGATTAGCCTCGCCTCT- 3077  
Db 3224 ATCCCTCCACCTATCTCTAACAGGCTGTAGATTACCAAAACAGATTTAATTCATCACTA 3283  
Qy 3078 -GAGGAAGCGCCTACAGCGGTTGCTTCGCTGGACTTTTCTCTAGATGCTGTGCTCAT 3136  
Db 3284 AAAGAAATCTATTATCAACTGCTGCTTCCACAGACTTTTCTCTAGAGCTGTCTGCGTT 3343  
Qy 3137 TACTC-----CAAGTGAATCTTATAAAATCAAAACC-TCTCCTGCACAGCGGGAGA 3188  
Db 3344 TACTCTGTTTCAAAGGAGCTTTTGAATAATCAATCTATCTCTCACAGGAGGAGGA 3403  
Qy 3189 GCAATAATGAGACTTTGTTGGTGGCCCGCTTACCCCTGGGGCTTTTCCACGAGCTTGAG 3248  
Db 3404 GTGATAATGAATCTTTATTTGGAGCATTTGATCTGCATCCAGGCTTCTCAGGCTGGCTG 3463  
Qy 3249 GGGAAAGCCTATCTATGAAATATAGTATATTCTTGTAAATACGTGAAACCAAAACC 3308  
Db 3464 AGTGA- TTGTGTACCTGAAGTACAGTATATCTTGTAAATACATAAAACAAA----- 3515  
Qy 3309 CGTTTTCCTAAGGAAAGCTAAATATGATTTTAAATCTATGTTTAAATATCTAT 3368  
Db 3516 AGCATTTTCTAAGGAGAGCTAAATGATTTT--AAGTCTATGTTTAAATATATAT 3573  
Qy 3369 GTAACTTTTCTATTTAGTATATATTTATGATGAGATGGAATAAATTTCTACTGTA 3428  
Db 3574 GTAAATTTTTCAGTATTTAGTATATATTTTATGGGTGGAAATAAATTTCTACTCAA 3633  
Qy 3429 AAAAAAATAAAAA 3442  
Db 3634 AAAAAAATAAAAA 3647

## RESULT 2

AK045865

LOCUS

DEFINITION

AK045865

Mus musculus adult male corpora quadrigenina CDNA, RIKEN

full-length enriched library, clone:B230315G04 product:FMS-like

tyrosine kinase 3, full insert sequence.

AK045865

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AK045865 3761 bp mRNA linear HTC 05-DEC-2002  
Mus musculus adult male corpora quadrigenina CDNA, RIKEN  
full-length enriched library, clone:B230315G04 product:FMS-like  
tyrosine kinase 3, full insert sequence.

AK045865.1 GI:26337656

HTC; CAP trapper.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1. Carninci, P. and Hayashizaki, Y.

High-efficiency full-length CDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

10349636



```
QY 481 GCCAACTACAGTACTGTTTCACAGTGAATGAAGAGATACACAGCTGTATGTCTAAGG 540
Db 498 GCCAACTACAGTACTGTTTCACAGTGAATGAAGAGATACACAGCTGTACGTGCTAAGA 557
QY 541 AGACCTTACTTTAGGAAGATGGAACACCAAGGATGCACTGCTCTGCACTCTCCGAGGGTGT 600
Db 558 AGACCTTACTTTAGGAAGATGGAACACCAAGGATGCACTGCTCTGCACTCTCCGAGGGTGT 617
QY 601 CCGAGGCCACCTGTGGAGTGGTGTCTGCACTCTCCAGTCCCAAGGAAAGCTGTAAAGAGAA 660
Db 618 CCAGAGGCCACCTGTGGAGTGGTGTCTGCACTCTCCAGTCCCAAGGAAAGCTGTAAAGAGAA 677
QY 661 GGCCTGTGTTGTTCAGAAAGGAGGAAAGGACTTTCATGAGTTGTTCCGACAGACATC 720
Db 678 GGCCTGTGTTGTTCAGAAAGGAGGAAAGGACTTTCATGAGTTGTTCCGACAGACATC 737
QY 721 AGATGCTGTGTAGAAATGCATCGGCCCGCGAATGACCAAGCTGTTCACCATAGATCTA 780
Db 738 AGATGCTGTGTAGAAATGCATCGGCCCGCGAATGACCAAGCTGTTCACCATAGATCTA 797
QY 781 AACAGGCTCTTCAGAGCACATGCCCCAGTTATTCTGAAAGTGGGGAAACCCTTGTGG 840
Db 798 AACAGGCTCTTCAGAGCACATGCCCCAGTTATTCTGAAAGTGGGGAAACCCTTGTGG 857
QY 841 ATCAGGTGAAGGCCATTCATGTGAACCATGATTCGGGCTCACCTGGGAGCTGGAAGAC 900
Db 858 ATCAGGTGAAGGCCATTCATGTGAACCATGATTCGGGCTCACCTGGGAGCTGGAAGAC 917
QY 901 AAAGCCCTGGAGGGGAGGAGTACTTTGAGATGAGTACTTCTCCACCAACAGGACCATG 960
Db 918 AAAGCCCTGGAGGGGAGGAGTACTTTGAGATGAGTACTTCTCCACCAACAGGACCATG 977
QY 961 ATTCCGATTCTCTTGCCCTTTGTCTCCGTTGGGAAGGAAACGACACCGGATATTACACC 1020
Db 978 ATTCCGATTCTCTTGCCCTTTGTCTCCGTTGGGAAGGAAACGACACCGGATATTACACC 1037
QY 1021 TGCTTCTCTCAAGACACCCGACAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 1080
Db 1038 TGCTTCTCTCAAGACACCCGACAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 1097
QY 1081 ATAAACGCTACAGCTCGAAGAGATGAATGAAATGACCCGTACGAAAGTCTCTCTTC 1140
Db 1098 ATAAACGCTACAGCTCGAAGAGATGAATGAAATGACCCGTACGAAAGTCTCTCTTC 1157
QY 1141 TCAGTCAGTGTAAAGCGTACCCACCAATCCGATGACGCTGGATCTTCTCTCAAGCCTCA 1200
Db 1158 TCAGTCAGTGTAAAGCGTACCCACCAATCCGATGACGCTGGATCTTCTCTCAAGCCTCA 1217
QY 1201 TTTCTTGTGAACAGAGGCGCTGGAGGATGGGTACAGCATATCTAAAT 1249
Db 1218 TTTCTTGTGAACAGAGGCGCTGGAGGATGGGTACAGCATCTCTCAAT 1266

RESULT 3
AI323253
LOCUS
DEFINITION
mp88b02.y1 Soares thymus 2N8MT Mus musculus cDNA clone IMAGE:576267
5', similar to gb:U02687.FL CYTOKINE RECEPTOR PRECURSOR (HUMAN);
gb:X59398 Mouse flt3 mRNA for tyrosine kinase receptor of the PDGF
(MOUSE);, mRNA sequence.
ACCESSION
AI323253
VERSION
AI323253.1 GI:4057682
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
REFERENCE
AUTHORS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 765)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Giesel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, B., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
```

```
Waterston, R.
The WashU-HMI Mouse EST Project
Unpublished
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:350915
This read is a RESEQUENCE of a previously sequenced mouse clone
correct orientation)
Seq primer: -40RP from Gibco
High quality sequence stop: 423.
FEATURES
Location/Qualifiers
source
1..765
/organism="Mus musculus"
/mol_type="mrna"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGS:576267"
/sex="male"
/tissue_type="Thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"
/clone_lib="Soares thymus 2N8MT"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTACCACTCAAGTGGGAGCGCGGTTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. RNA
provided by Dr. Bertrand Jordan. Library went through two
rounds of normalization, and was constructed by Bento
Soares and M.Fatima Bonaldo."
BASE COUNT 203 a 176 c 182 g 200 t 4 others
ORIGIN
Query Match 19.0%; Score 655.8; DB 9; Length 765;
Best Local Similarity 94.6%; Pred. No. 3.2e-74;
Matches 699; Conservative 0; Mismatches 36; Indels 4; Gaps 2;
QY 1558 GGGCTTCTGGTCAAATGCTGGTACAAATCTTATGGGACGCTTCGGAACCATCTTT 1617
Db 1 GGGCTTCTGGTCAAATGCTGGTACAAATCTTATGGGACGCTTCGGAACCATCTTT 60
QY 1618 TTAAACTCACCAGGCCCTTCCCTTTCATCCAGACAAATCTCTTATGGACCATTT 1677
Db 61 TTAAACTCACCAGGCCCTTCCCTTTCATCCAGACAAATCTCTTATGGACCATTT 120
QY 1678 GGGCTTCTGCCCTTCATTTGTTTCTCATTTGTTGATCTGCCCAAAATCAAAAAG 1737
Db 121 GGGCTTCTGCCCTTTCATTTGTTTCTCATTTGTTGATCTGCCCAAAATCAAAAAG 180
QY 1738 CAATTAGTACGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1797
Db 181 CAATTAGTACGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 240
QY 1798 TACTTCTAGCTTGACTTCAGGGACTATGAATGACCTTAAGTGGGAGTTCCCGAGAGAG 1857
Db 241 TACTTCTAGCTTGACTTCAGGGACTATGAATGACCTTAAGTGGGAGTTCCCGAGAGAG 300
QY 1858 AACTTAGAGTTTGGGAAGGTCTCGGGCTCTGGCGCTTTTCGGGAGGGTGAATGAACGCCACG 1917
Db 301 AACTTAGAGTTTGGGAAGGTCTCGGGCTCTGGCGCTTTTCGGGAGGGTGAATGAACGCCACG 360
QY 1918 GCCTATGGCATTAGTAAACCGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTTAAAGAG 1977
Db 1918 GCCTATGGCATTAGTAAACCGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTTAAAGAG 1977
```





discovery  
Genome Res. 6 (9), 791-806 (1996)  
9704477  
8889548  
Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@uiowa.edu  
cDNA Library Preparation: M.B. Soares Lab Clone distribution:  
clones will be available through Research Genetics (www.resgen.com)  
This clone is also available through the I.M.A.G.E. Consortium at  
LLNL (info@image.llnl.gov). IMAGE ID= 1793119  
Seq primer: M13 Forward.

# FEATURES

source  
1. .604  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/clone="UI-R-C2p-rg-f-10-0-UI"  
/dev\_stage="adult"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="UI-R-C2p"  
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-C2p library is a subtracted library derived from the UI-R-C1 library, which is a subtracted library derived from the UI-R-C0 library. The UI-R-C0 library consisted of a mixture of individually tagged normalized libraries constructed from rat placenta, adult lung, brain, liver, kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day embryo. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-dT track which allows identification of the library of origin of a clone within the mixture. The subtracted library (UI-R-C2p) was constructed as follows: PCR amplified cDNA inserts from UI-R-C1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with the UI-R-C1 library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-C2p library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996)."

BASE COUNT 166 a 132 c 167 g 139 t  
ORIGIN  
Query Match 15.4%; Score 531.8; DB 10; Length 604;  
Best Local Similarity 93.7%; Pred. No. 2e-58;  
Matches 554; Conservative 0; Mismatches 37; Indels 0; Gaps 0;  
QY 1991 GTGAAAAGAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGACCATGACA 2050  
DB 8 GTGAAGAGAGGGCTCATGCTGAGCTCAAAATGATGACCCACCTGGGACCATGACA 67  
QY 2051 ACATCGTGAATCTGCTGGGGGCAATGACA CATGTCAGGGCCAGTGTACTTGAATTTTGAAT 2110  
DB 68 ACATCGTGAACCTGCTGGGGGCAATGACA CATGTCAGGGCCAGTGTACTTGAATTTTGAAT 127  
QY 2111 ATTGTTGCTATGCTGACCTCTCAACTACCTAAGAGTAAAGAGAGAGTTTCACAGGA 2170  
DB 128 ATTGTTGCCATGCTGACCTCTCAACTACCTAAGAGTAAAGAGAGAGTTTCACAGGA 187  
QY 2171 CATGGAAGAGATTTTAAAGAAATAATTTTCACTTCTTACCCTTCTTCCAGGCACATT 2230  
DB 188 CGTGGACAGAGATTTTAAAGAAATAATTTTCACTTCTTACCCTTCTTCCAGGCACATT 247  
QY 2231 CAAATCCAGCATGCTGCTGTTTACGAGAGTTTCACTTCTTACCCTTCTTCCAGGCACATT 2290

Db 248 CAAACTCAAGTATGCCGGTTTCACAGAGTTTCAGATATACCGCCCGTGTGATCAGGTCT 307  
QY 2291 CAGGGTTCAATGGGAATTCAATTCAATTCTGAAGATGAGATTGAATATGAAAACCAAGAAGA 2350  
Db 308 CAGGGTTCAATGGGAATTCAATTCAATTCTGAAGATGAGATTGAATATGAAAACCAAGAAGA 367  
QY 2351 GGCTGGCAGAGAAGAGAGAGAGAGATTGAAACGTGCTGACGCTTTGAAAGACCTCTTGTCT 2410  
Db 368 GGCTGGAAGAGAGAGAGAGAGATTGAAACGTGCTGACGCTTTGAAAGACCTCTTGTCT 427  
QY 2411 TTGGTACCAGTGGCCAAAGGAGTGAATTCCTGGAGTTCAAGTCGCTGTCACACAG 2470  
Db 428 TTGGTATCAAGTGGCCAAAGGAGTGAATTCCTGGAGTTCAAGTCGCTGTCACACAG 487  
QY 2471 ACCTGGCAGCAGGAGATGTGTGTGTGTCACCCACGGGAGGCTGCTGAAGATCTGTGACTTTG 2530  
Db 488 ACCTGGCAGCAGGAGATGTGTGTGTGTCACCCACGGGAGGCTGCTGAAGATCTGTGACTTTG 547  
QY 2531 GACTGGCCCGAGACATCTGAGCGACTCCAGCTACGTCGTCAGGGGCAACG 2581  
Db 548 GACTGGCCCGAGACATCTGAGTCACTCCAGTACGTCCTCAGGGGCAACG 598

# RESULT 6

AI323643/c  
LOCUS  
DEFINITION  
m88b02.x1 Soares thymus 2NBMT Mus musculus cDNA clone IMAGE:576267  
3', similar to gb:U02687 FL CYTOKINE RECEPTOR PRECURSOR (HUMAN);  
gb:XS9398 Mouse Flt3 mRNA for tyrosine kinase receptor of the PDGF  
(MOUSE);, mRNA sequence.  
ACCESSION  
AI323643  
VERSION  
AI323643.1 GI:4058072  
KEYWORDS  
EST.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
REFERENCE  
1 (bases 1 to 747)  
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,  
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,  
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and  
Waterston, R.  
TITLE  
The WashU-HMI Mouse EST Project  
JOURNAL  
Unpublished  
COMMENT  
Contact: Marra M/Mouse EST Project  
WashU-HMI Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:350915  
This clone was previously sequenced on the 5' end only, this new  
data is from the 3' end  
High quality sequence stop: 432.

# FEATURES

source  
1. .747  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:576267"  
/sex="male"  
/tissue\_type="Thymus"  
/dev\_stage="4 weeks"  
/lab\_host="DH10B"  
/clone\_lib="Soares thymus 2NBMT"  
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA





```

RESULT 9
BI360262
LOCUS      387099 MARC 2P1G Sus scrofa cDNA 5', mRNA linear EST 01-AUG-2001
DEFINITION
ACCESSION  BI360262
VERSION     BI360262.1 GI:15056290
KEYWORDS   EST.
SOURCE     Sus scrofa (pig)
ORGANISM   Sus scrofa
REFERENCE  1 (bases 1 to 621)
AUTHORS   Fahrenkrug, S.C., Smith, T.P.L., Freking, B.A., Cho, J., White, J.,
          Vallet, J., Wise, T., Rohrer, G.A., Perte, G., Sultana, R., Quackenbush
          , J. and Keele, J.W.
TITLE     Porcine gene discovery by normalized cDNA-library sequencing and
          EST cluster assembly
JOURNAL   Mamm. Genome 13 (8), 475-478 (2002)
MEDLINE   22213789
PUBMED    12228715
COMMENT   Contact: Smith TPL
          USDA, ARS, US Meat Animal Research Center
          PO Box 166, Clay Center, NE 68933-0166, USA
          Tel: 402 762 4366
          Fax: 402 762 4390
          Email: smith@email.marc.usda.gov
          Single pass sequencing. Bases called and alt trimmed with phred
          v0.980904.e. Vector identified by cross_match with the -minscore 18
          and -mismatch 12 options.
          PCR Primers
          FORWARD: AGGAACAGCTATGACCAT
          BACKWARD: GTTTCGCCAGCTACGACG
          Plate: 139 row: H column: 2
          Seq primer: ATTTAGGTGACACTATAG.
FEATURES
source     Location/Qualifiers
            1..621
             /organism="Sus scrofa"
             /mol_type="mRNA"
             /db_xref="taxon:9823"
             /tissue_type="pooled"
             /lab_host="DH10B"
             /clone_lib="MARC 2P1G"
             /notes="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
             Library made from pooled tissue from testis, ovary,
             endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT 190 a 132 c 141 g 158 t
ORIGIN
Query Match      13.1%; Score 452.6; DB 12; Length 621;
Best Local Similarity 83.2%; Pred. No. 2.4e-48;
Matches 515; Conservative 0; Mismatches 104; Indels 0; Gaps 0;
QY 1702 GTTCTCATGTGTGTGTTCTCCCAAAATACAAAAGCAATTTAGGTACGAGTCTAGCTG 1761
Db 3 GTTTTAAACGATGTAATTTGTCAAGTACAAAAGCGATTCGGTACGAAAGCCAGCTG 62
QY 1762 CAGATGATCCAGTGTACTGCCGCCCTCGGATAACGAGTACTTCTACGTTGACTTCAGGGAC 1821
Db 63 CAGATGTCAGGTGACGGGGTCCCTGGATAACGACTTCTTACATCGACTTCAGAGAA 122
QY 1822 TATGAATATGACCTTAAGTGGGAGTTCGCCGAGAGAGAACTTAGAGTTGGGAGGCTCTG 1881
Db 123 TATGAATATGACCTCAATGGGAGTTTCCGAGGAAAAATTTAGAAATTTGGGAGGTTCTG 182
QY 1882 GGCTCTGGCGCTTTCGGGGGGTGTGATGACGCCAGCGCTATGCGATTACTTAAACGGGA 1941
Db 183 GGATCGCGCGCTTTGGAAAGGTGATGACGCACTGCCTTATGGAATCAGTAAACCTGGA 242
QY 1942 GTCTCAATTACAGTGGCGGTGAAGATGCTTAAAGAGAGAAAGCTGACAGCTGTGAAAAAGAA 2001
Db 243 GTGTCAATCCAGTGGCAGTCAAAATGCTGAAGAAAGGACGACAGCTCGGACGAGAG 302
QY 2002 GCTCTCATGTGGAGCTCAAAATGATGATGACCCACCTGGGACCCATGACAAACATCGTGAAT 2061

```

```

Db 303 GCTCTCATGTGTGAATCAAAATGATGACCCACCTGGGCGAGCCAGAGATATAGTGAAC 362
QY 2062 CTGCTGGGGCGATGCACACATGTCTCAGGGCCAGTGTACTTGTGATTTTGAATATTTGTGCTAT 2121
Db 363 CTGCTGGGGCGCTGCACCCCTGTCTCAGGGCCCATTTACTTGTGATTTTGAATATTTGTGCTAC 422
QY 2122 GGTGACCTCTCAACTACCTAAAGAGTAAAGAGAGAGAGTTCACAGGACATGACAGAG 2181
Db 423 GGTGATCTTCTCAACTATCTAAAGAGTAAAGAGAGAGAAATTTCTAGGACATGACCGAG 482
QY 2182 ATTTTAAAGAACATATTTTCACTTCTTACCTACTTTCAGGCGACATTCAAATCCAGC 2241
Db 483 ATTTTCAAGGAACATATTTTCACTTCTTATCTTCTTATCTTCCAAATCACACCCAAATTCAGT 542
QY 2242 ATGCTCTGTTTTCAGGAGAGTTCAGTTACACCGCCCTTGGATCAGCTCTCAGGGTTCAAT 2301
Db 543 ATGCGGGTTCAGAGAGTTCAAATACACCCAGCTCGGATCTCTATCTCAGGATTCAT 602
QY 2302 GGGAAATTCATTCATTCG 2320
Db 603 GGGAAATTCATTCATTCG 621

RESULT 10
BY727686 658 bp mRNA linear EST 17-DEC-2002
LOCUS      BY727686 RIKEN full-length enriched, 6 days neonate medulla
DEFINITION oblongata Mus musculus cDNA clone B73004D22 5', mRNA sequence.
ACCESSION  BY727686
VERSION     BY727686.1 GI:27140813
KEYWORDS   EST.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
REFERENCE  1 (bases 1 to 658)
AUTHORS   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
          Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
          Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H.,
          Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,
          Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A.,
          Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S.,
          Beissel, K.W., Blake, J.A., Bradt, D., Bruscia, V., Chothia, C., Corbani
          , L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest
          , A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.
          , Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J.,
          Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M.,
          King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons
          , P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki
          , H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Perte, G.,
          Pesole, G., Petrowsky, N., Pillai, R., Pontius, J.U., Qi, D.,
          Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring
          , B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semplic, C.A., Secou
          , M., Shmida, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale
          , R.D., Tomita, M., Verard, R., Wagner, L.G., Wynshaw-Boris, A., Yanagisawa
          , Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa
          , M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.
          , Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura
          , M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.
          , Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Iehi
          , Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata
          , K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander
          , E.S., Rogers, J., Birney, E. and Hayashizaki, Y.
          Analysis of the mouse transcriptome based on functional annotation
          of 60,770 full-length cDNAs
          Nature 420, 563-573 (2002)
JOURNAL    22354683
MEDLINE    12466851
PUBMED     12466851
COMMENT    Contact: Yoshihide Hayashizaki
          Laboratory for Genome Exploration Research Group, RIKEN Genomic
          Sciences Center (GSC), Yokohama Institute
          The Institute of Physical and Chemical Research (RIKEN)
          1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

```

Qy	121	CTGCCTGTGATCAAGTGTGTTTTAATCAATCATGGAACAATGGCTTCATCAGCGGAAAG	180
Db <td>316</td> <td>CTGCCTGTGATCAAGTGTGTTTTAATCAATCATGGAACAATGGCTTCATCAGCGGAAAG</td> <td>375</td>	316	CTGCCTGTGATCAAGTGTGTTTTAATCAATCATGGAACAATGGCTTCATCAGCGGAAAG	375
Qy <td>181</td> <td>CCATCATCGTACCGAATGGTGCAGGATCCCCCAGAGAGCCTCCAGTGTACCCCGAGCGC</td> <td>240</td>	181	CCATCATCGTACCGAATGGTGCAGGATCCCCCAGAGAGCCTCCAGTGTACCCCGAGCGC	240
Db <td>376</td> <td>CCATCATCGTACCGAATGGTGCAGGATCCCCCAGAGAGCCTCCAGTGTGCCCCGAGCGC</td> <td>435</td>	376	CCATCATCGTACCGAATGGTGCAGGATCCCCCAGAGAGCCTCCAGTGTGCCCCGAGCGC	435
Qy <td>241</td> <td>CAGAGTGAAGGGACGGTATATGAAGCGGCCACCGTCGGAGGTGSCCGAGTCTGGGTCCATC</td> <td>300</td>	241	CAGAGTGAAGGGACGGTATATGAAGCGGCCACCGTCGGAGGTGSCCGAGTCTGGGTCCATC	300
Db <td>436</td> <td>CAGAGTGAAGGGACGGTATATGAAGCGGCCACCGTCGGAGGTGSCCGAGTCTGGGTCCATC</td> <td>495</td>	436	CAGAGTGAAGGGACGGTATATGAAGCGGCCACCGTCGGAGGTGSCCGAGTCTGGGTCCATC	495
Qy <td>301</td> <td>ACCTCTCAAGTGCAGTCTGCGCACCCAGGGGACCTTTCTGCTCTGGGTCTTTAAGCAC</td> <td>360</td>	301	ACCTCTCAAGTGCAGTCTGCGCACCCAGGGGACCTTTCTGCTCTGGGTCTTTAAGCAC	360
Db <td>496</td> <td>ACCTCTCAAGTGCAGTCTGCGCACCCAGGGGACCTTTCTGCTCTGGGTCTTTAAGCAC</td> <td>555</td>	496	ACCTCTCAAGTGCAGTCTGCGCACCCAGGGGACCTTTCTGCTCTGGGTCTTTAAGCAC	555
Qy <td>361</td> <td>AGTCTCCTGGGTGCGAGCGGCATTTGATTTACAAAACAGAGGAATCGTTTCCATGGCC</td> <td>420</td>	361	AGTCTCCTGGGTGCGAGCGGCATTTGATTTACAAAACAGAGGAATCGTTTCCATGGCC	420
Db <td>556</td> <td>AGTCTCCTGGGTGCGAGCGGCATTTGATTTACAAAACAGAGGAATCGTTTCCATGGCC</td> <td>615</td>	556	AGTCTCCTGGGTGCGAGCGGCATTTGATTTACAAAACAGAGGAATCGTTTCCATGGCC	615
Qy <td>421</td> <td>ATCTTGAAAGTGCACAGACCCAGGCGAGGAGATCCTACTCC</td> <td>463</td>	421	ATCTTGAAAGTGCACAGACCCAGGCGAGGAGATCCTACTCC	463
Db <td>616</td> <td>CATCTTGAGTGCACAGACCCAGGCGAGGAGATCCTACTCC</td> <td>658</td>	616	CATCTTGAGTGCACAGACCCAGGCGAGGAGATCCTACTCC	658

RESULT 11

BE851447/c

LOCUS

DEFINITION

445 bp mRNA linear EST 06-SEP-2000

BE851447

UI-M-BH3-atk-d-06-0-01.1 NIH BMAP M.S4 Mus musculus cDNA clone

UI-M-BH3-atk-d-06-0-01.5 mRNA seq2000

ACCESSION BE651447  
 VERSION BE651447.1 GI:9977271  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 445)  
 REFERENCE Ronaldo,M.F., Lennon,G. and Soares,M.B.  
 AUTHORS Normalization and subtraction: two approaches to facilitate gene  
 TITLE discovery

discovery  
Genome Res. 6 (9), 791-806 (1996)  
97044477  
8889548  
Contact: Chin, H  
National Institute of Mental Health  
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
20892-9643, USA  
Tel: 301 443 1706  
Fax: 301 443 9890  
Email: mEST@mail.nih.gov  
CDNA Library Preparation: M.B. Soares Lab Clone distribution:  
Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It  
should be noted that Bento Soares is generating a small number of  
additional specialized non-redundant arrays of BMAP cDNAs whose  
availability will be considered under appropriate and limited  
collaborative arrangements  
corresponding

```

FEATURES
  source
seq primer: MUS reverse.
  Location/Qualifiers
    1..445
      /organism="Mus musculus"
      /mol_type="mRNA"
      /strain="C57BL/6J"
      /db_xref="taxon:10090"
      /clone="UI-M-BH3-ak-d-06-0-UI"
      /dev_stage="27-32 days"
      /lab_host="DH10B (Life Technologies)"
      /clone_lib="NIH BMAP M 54"
      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site1: Not I; Site 2: Eco RI; The
NIH_BMAP_M_54 library is a subtracted library of a series,
ultimately derived from a mixture of individually tagged

```

(cerebellum, brain stems, olfactory bulbs, hypothalamus, cortex, amygdala, basal ganglia, pineal gland, striatum, hippocampus) after a series of subtractions to reduce the representation of cDNAs from which ESTs had already been generated. The following serially subtracted libraries were generated in this process: NIH BMAP M S4, NIH BMAP M S3.3, NIH BMAP M S3.2, NIH BMAP M S3.1, NIH BMAP M S2, NIH BMAP M S1. The subtracted library (NIH BMAP M S4) was constructed as follows: PCR amplified cDNA inserts from NIH BMAP M S3.3, NIH BMAP M S3.2, and NIH BMAP M S3.1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with a pool of the NIH BMAP M S3.3, NIH BMAP M S3.2, and NIH BMAP M S3.1 libraries in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the NIH BMAP M S4 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)

BASE COUNT 96 a 123 c 111 g 115 t

Query Match 12.9%; Score 445; DB 10; Length 445;  
Best Local Similarity 100.0%; Pred. No. 2.6e-47;  
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 612 TGTGAGTGGGCTCTGAGCTCCACAGGAAAGCTGTAAGAGAGGCGCTGCTGT 671  
Db |||||  
QY 445 TGTGAGTGGGCTCTGAGCTCCACAGGAAAGCTGTAAGAGAGGCGCTGCTGT 386  
Db |||||  
QY 672 TGTGAGTGGGCTCTGAGCTCCACAGGAAAGCTGTAAGAGAGGCGCTGCTGT 731  
Db |||||  
QY 385 TGTGAGTGGGCTCTGAGCTCCACAGGAAAGCTGTAAGAGAGGCGCTGCTGT 326  
QY 732 TAGAATGCACTGGGCGGGAATGACCAAGCTGTTACCATAGATCTAAACAGAGCTCC 791  
Db |||||  
QY 325 TAGAATGCACTGGGCGGGAATGACCAAGCTGTTACCATAGATCTAAACAGAGCTCC 266  
QY 792 TCAGAGCACTGCGCCAGTATTCTGAAAGTGGGGAACCTTGTGGATCAGGTGTA 851  
Db |||||  
QY 265 TCAGAGCACTGCGCCAGTATTCTGAAAGTGGGGAACCTTGTGGATCAGGTGTA 206  
QY 852 GGCCATCCATGTGAACCATGATTCGGCTCACCTGGAGCTGGAAGCAAAAGCCCTGGA 911  
Db |||||  
QY 205 GGCCATCCATGTGAACCATGATTCGGCTCACCTGGAGCTGGAAGCAAAAGCCCTGGA 146  
QY 912 GGAGGCGAGCTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATGATTCGGATTCT 971  
Db |||||  
QY 145 GGAGGCGAGCTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATGATTCGGATTCT 86  
QY 972 CTGTCCTTTGCTCTCCGTTGGGGAAGCAACACACCGGATATTACCTGCTCTTCCTC 1031  
Db |||||  
QY 85 CTGTCCTTTGCTCTCCGTTGGGGAAGCAACACACCGGATATTACCTGCTCTTCCTC 26  
QY 1032 AAAGCACCCCGCAGTCCAGCTTG 1056  
Db |||||  
QY 25 AAAGCACCCCGCAGTCCAGCTTG 1

RESULT 12  
BI461248  
LOCUS 714 bp mRNA linear EST 21-AUG-2001  
DEFINITION 603206574F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:5272266 5',  
mRNA sequence.  
BI461248  
ACCESSION BI461248.1 GI:15251904  
VERSION EST.  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 (bases 1 to 714)  
NIH-MGC http://mgc.mci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D. (NHGRI), Shiraki  
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
Toshiyuki and Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: InCyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LAM11687 row: c column: 19  
High quality sequence stop: 709.

FEATURES  
source

1..714  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5272266"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_97"  
/note="Organ: testis; Vector: pBluescriptR (modified  
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag  
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTN-3',  
size-selected for average insert size 2.2 kb and  
normalized to ROT 5. This is a primary library enriched  
for full-length clones and constructed using the  
Cap-trapper method (Carninci, in preparation). Library  
constructed by M. Brownstein (NIH/NHGRI, National  
Institutes of Health). Note: this is a NIH\_MGC Library."

BASE COUNT 182 a 171 c 191 g 170 t

Query Match 12.5%; Score 432.8; DB 12; Length 714;  
Best Local Similarity 79.5%; Pred. No. 7.2e-46;  
Matches 525; Conservative 0; Mismatches 132; Indels 3; Gaps 1;

QY 1 GCGGCTGGCTACCGCGCTCCGAGGCGCATCGGGGTTGGCGGAGCGGCGGCGG 60  
Db |||||  
QY 58 GCGGCTGGGAGCGCGCGGCTCCGAGGCGCATCGCGGCGTGGC--GCGGCGCGCGC 114  
QY 61 CGGCTGCTGCTGCTGTTGTTTGTGCTAATGATTTCTTGACACCGTTTACAAACCAAGAC 120  
Db |||||  
QY 115 CAGCTGCCGCTGCTGCTGTTTCTGCAATGATATTGGGACTATTACAATCAAGAT 174  
QY 121 CTGCTGTGATCAAGTGTGTTTTTAATCATGATGAGAAATGGGTCTATCAGCGGGAAG 180  
Db |||||  
QY 175 CTGCTGTGATCAAGTGTGTTTTTAATCAATCAATGATGATTCATCAGTGGGGAAG 234  
QY 191 CCATCATGTCACCGAATGGTGGGAGTCCCGAGAGACCTCAGTGTACCCGAGGCGC 240  
Db |||||  
QY 235 TCATCATCATATCCCATGTTATCAGAAATCCCGGGAAGACCTCGGCTGTCGTTGAGACC 294  
QY 241 CAGAGTGAAGGAGCGGTATATGAAGCGCCACCGTGGAGTGGCCGAGTCTGGGTCCATC 300  
Db |||||  
QY 295 CAGAGTCAAGGAGCAGTGTGTAAGAGCTCCGCTGTTGGAAGTGTATCTGCTTCCATC 354  
QY 301 ACCCTGCAAGTGCAGCTCGCCACCCGAGGAGACCTTCTGCTCTGCGTCTTTAAGCAC 360  
Db |||||  
QY 355 ACCTGCAAGTGTGCTGATGTCGATGTCGCGGAGGAAACATTTCTGCTCTTTAAGCAC 414  
QY 361 AGCTCCCTGGGCTGCCAGCGGAGCTTTGATTTTACAAACAGAGGAATCGTTCCATGGCC 420  
Db |||||  
QY 415 AGCTCCCTGAAATGCCAGCCCATTTTGAATTTACAAACAGAGGAGTAGTTTCCATGTC 474  
QY 421 ATCTTGAACGTCAGAGAGCCAGGAGAGTAACCTACTCCATATTCAGAGCGGAGCC 480  
Db |||||  
QY 475 ATTTGAAATGACAGAAACCCAGCTGAGAAATACCTACTTTTATTCAGAGTGAAGCT 534  
QY 481 GCCAACTACAGTACTGTTTCAAGTGAATGTGAAGAGATACACAGCTGTATGTGCTAAGG 540



db	535	ACCAATTTACACNAATTATTGTTTTCAGTGTAGTATAGAAATACCTGCTTTACACATTAGA	594
Qy	541	AGACCTTTACTTTAGGAAGATGAAAAACGAGATGCATCTGCTCTGCAATCTCCGAGGGTGT	600
db	595	AGACCTTTACTTTAGAAAAATGAAAAACGAGAGCCCTGGTCTGCAATCTGAGAGCGTT	654
Qy	601	CCGAGAGCCCACTGTGGAGTGGTGCTCTGCAAGTCCCAACGGAAGCTGTAAAGAGAAGA	660
db	655	CCAGAGCCGATCGTGTGAAATGGTGCTTTTGCATTCACAGGGGGAAGCTGTAAAGAGAACA	714

RESULT	13
BU695104	
LOCUS	
DEFINITION	BU695104 597 bp mRNA linear EST 08-OCT-2002 HLI2in11905T7 Hematopoietic Stem Cell Subtracted Library Mus musculus cDNA 5' similar to ftk-2, mRNA sequence.

ACCESSION	BU695104
VERSION	BU695104.1
KEYWORDS	GI:23596587
SOURCE	EST.
ORGANISM	Mus musculus (house mouse)
REFERENCE	Mus musculus
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 597)
TITLE	Phillips,R.L.; Ernst,R.E.; Brunk,B.P.; Ivanova,N.; Mahan,M.A., Deanehan,J.K.; Moore,K.A.; Overton,G.C. and Lemischka,I.R. The genetic program of hematopoietic stem cells
JOURNAL	Science 288 (5471), 1635-1640 (2000)
MEDLINE	20295303
PUBLISHED	10834841

CONTACT: Lemischka, Ihor R.  
Department of Molecular Biology  
Princeton University  
Lewis Thomas Laboratory, Washington Road, Princeton, NJ 08544, USA  
Tel: 609 258 2838  
Fax: 609 258 2759  
Email: [illemischka@molbio.princeton.edu](mailto:illemischka@molbio.princeton.edu)  
These ESTs are derived from a subtracted cDNA library enriched for  
gene products expressed in day 14-14.5 fetal liver hematopoietic  
stem cells defined as Lineageneg/lo, AA4.lpos, cKitpos, Ly6A/E  
(Sca-1)pos

```

seq primer: M13Reverse or T7.
Location/Qualifiers
1..597
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57Bl/6J"
/db_xref="taxon:10090"
/tissue_type="Hematopoietic"
/cell_type="Stem Cells; Lineageneg/lo, AA4.lpos, cKitpos,
Ly6A/E (Sca-1)pos"
/dev_stage="Embryonic day 14-14.5"
/lab_host="DH10B"
/clone_lib="Hematopoietic Stem Cell Subtracted Library"
/note="Organ: Fetal Liver; Vector: Sport 1; Site 1: Sal I
Site 2: Not I; Two directionally cloned cDNA libraries
were made from fetal liver hematopoietic stem cells
enriched to be Lineageneg/lo, AA4.lpos, cKitpos, Ly6A/E
(Sca-1)pos called Scapos and from AA4.lneg fetal liver
cells. Subtractive hybridization was performed by
hybridization of the target, Scapos, single stranded cDNA
library in pSport1 to biotinylated RNA transcribed from
the driver, AA4.lneg cDNA library in pSport2 with insert
cloned in the complementary orientation. For detailed
protocols and additional information please see our
website at http://stemcell.princeton.edu."
136 a 140 c 152 g 131 t 38 others
BASE COUNT
ORIGIN
Query Match 12.4%; Score 428.4; DB 13; Length 597;
Best Local Similarity 92.3%; Pred. No. 2.9e-45;
Matches 469: Conservative 0; Mismatches 35; Indels 4; Gaps

```

Query Match	12.4%	Score 428.4;	DB 13;	Length 597;
Best Local Similarity	92.3%	Pred. No. 2.9e-45;		
Matches 469;	Conservative	0; Mismatches 35;	Indels 4;	Gaps 3

724	Qy	TGCTGTGCTAGAAA	TGCAC	TGGCGCGCGCAATGCACCAAC	CTGTTCACCATAGATCTAAAC	783
2	Db	TGCTGTGCTAGANA	TGCAC	TGGCGCGCGCAATGCACCAAC	CTGTTCACCATAGATCTAAAC	61
784	Qy	CAGGCTCTCTCAGAGCACA	CTGCCCCCAGCTTTATTC	TCTGAAAGTGGGGAA	CCCTTGTGGATC	843
62	Db	CAGGCTCTCTCAGAGCACA	CTGCCCCCAGCTTTATTC	TCTGAAAGTGGGGAA	CCCTTGTGGATC	121
844	Qy	AGGTGTAAAGGCAT	CCATGTGAACCA	TGATTCGGGCTCACCTGGGAGCTGGAAGACAAA	903	
122	Db	AGGTGTAAAGGCAT	CCATGTGAACCA	TGATTCGGGCTCACCTGGGAGCTGGAAGACAAA	181	
904	Qy	GCCTCTGGAGGAGGCA	GCTACTTTTGAGATGAGTAC	TACTCTCCACAAACAGGACCATGAT	963	
182	Db	GCCTCTGGAGGAGGCA	GCTACTTTTGAGATGAGTAC	TACTCTCCACAAACAGGACCATGAT	241	
964	Qy	CGGATTCCTCTTGGCCTTTTGTGTCTTC	CTGGGGAAGAA	CGACACGGATATTACACTGC	1023	
242	Db	CGGATTCCTCTTGGCCTTTTGTGTCTTC	CTGGGGAAGAA	CGACACGGATATTACACTGC	301	
1024	Qy	TCTTCTCTAAAGCACA	CCCCAGCCAGTCACGCGT	TGGTGACCATCTCTAGAAAAGGGTTTATA	1083	
302	Db	TCTTCTCTAAAGCACA	CCCCAGCCAGTCACGCGT	TGGTGACCATCTCTAGAAAAGGGTTTATA	361	
1084	Qy	AACGCTACCAAGCT	CGCAAGAAGATATGAAAT	TGACCCGCTACGAAAGTTCTGCTTCTCA	1143	
362	Db	AACGCTATCAGNT	CGCAAGANGATNTGAAAT	TGACCCGCTACGAAAGTTCTGCTTCTCA	421	
1144	Qy	GTCAAGTTT	-AAGCGTACCACGAATCCGATCCAGTCG	-ATCTTCTCTCAAGCCTCA	1200	
422	Db	GTCCGGTTTAAAGCGTAC	CAAGGAATCCGATCNCNGGGATCTT	TCTCTNAAAGCCTCA	481	
1201	Qy	-TTTCTCTGTGAACAGAGAGCGCTGGAG	1227			
482	Db	TTTTTCTCTGTGANNNGANAGCGCTCGG	509			

RESULT 14	515 bp	linear	EST 20-JUN-2002
BQ556177		mRNA	
LOCUS	BQ556177-5	NIA Mouse 7.4K cDNA Clone Set	Mus musculus cDNA clone.
DEFINITION	H4038E08-5'	mRNA sequence.	

ACCESSION	B0556177
VERSION	B0556177.1
KEYWORDS	GI:21457065
SOURCE	EST.
ORGANISM	Mus musculus (house mouse)
REFERENCE	Mus musculus
AUTHORS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
TITLE	1 (bases 1 to 515) VanBuren,V., Piao,X., Dudekula,D.B., Qian,Y., Carter,M.G., Martin P.R., Stagg,C.A., Bassey,U., Alba,K., Hamatani,T., Kargul,G.J., Luo,A.G., Kelso,J., Hide,W. and Ko,M.S.H. Assembly, verification, and initial annotation of NIA 7.4K mouse cDNA clone set
JOURNAL	Genome Res. 12 (12), 1999-2003 (2002)
MEDLINE	22354164
PUBMED	12466305
COMMENT	Other ESTs: H4038E08-3

Contact: Yong Qian  
Laboratory of Genetics  
National Institute on Aging/National Institutes of Health  
333 Cassell Drive, Suite 3000, Baltimore, MD 21224-6820, USA  
Email: cana@geun.grc.nia.nih.gov  
This clone set has been freely distributed to the community. Please  
visit [http://19sun.grc.nia.nih.gov/cDNA/NIA\\_7\\_4k.html](http://19sun.grc.nia.nih.gov/cDNA/NIA_7_4k.html) for details.  
Plate: H4038 row: E column: 08  
Seq primer: -21M13 Reverse  
High quality sequence stop: 515  
POLYA=No.

FEATURES	
source	Location/Qualifiers
1. .515	

```
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="niaEST:H4038E08-5"
/clone="H4038E08"
/sex="mixed"
/dev_stage="mixed"
/lab_host="DH10B"
/notes="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This clone is among a rearrayed set of 7,407 clones from more than 20 cDNA libraries."

BASE COUNT      139 a  138 c  128 g  110 t
ORIGIN

Query Match      12.3%; Score 426.2; DB 13; Length 515;
Best Local Similarity 92.8%; Pred. No. 5.9e-45;
Matches 478; Conservative 0; Mismatches 13; Indels 24; Gaps 2;

QY 2725 GCTAACTTCTATAAAGTATTGAGTGGATTTAAATGGAGCAGCATTTCTATGCCACA 2784
      |||||
Db 1 GCTAACTTCTATAAAGTATTGAGTGGATTTAAATGGAGCAGCATTTCTATGCCACA 60

QY 2785 GAAGGATATATCTTCTAATGCAATCTGCTGGGCTTTTGACTCAAGGAAGCGCCATCC 2844
      |||||
Db 61 GAAGGATATATCTTCTAATGCAATCTGCTGGGCTTTTGACTCAAGGAAGCGCCATCC 120

QY 2845 TTCCCCAAGCTGACTTTCATTTTGGATGTGCTGGCAGGAGCAGAGAAGC----- 2897
      |||||
Db 121 TTCCCCAAGCTGACTTTCATTTTGGATGTGCTGGCAGGAGCAGAGAAGCGATGAT 180

QY 2898 -----ATGATCAGAAATCATTCATCTACCAATCTACCAAGAGCGGCGCCCTC 2941
      |||||
Db 181 CAGAAATCATGGTGGCAAGCTCCAGAAATCATTCATCTACCAAGAGCGGCGCCCTC 240

QY 2942 ACAGAGAGCGGGCTCAGAGCC-CAGTCCACAGCGCCAGGTGAAGTTCACAGAA 3000
      |||||
Db 241 ACAGAGAGCGGGCTCAGAGCCCGCCATCGCCACAGCGCCAGGTGAAGTTCACAGAA 300

QY 3001 AGAAGTTAGCGAGGAGCGCTTGACCCCGCCCTAGCAGGCTGTAGACCGCAGAGCCA 3060
      |||||
Db 301 AGAAGTTAGCGAGGAGCGCTTGACCCCGCCCTAGCAGGCTGTAGACCGCAGAGCCA 360

QY 3061 AGATTAGCTCGCTCTGAGGAAGCGCCCTACAGCGGTTGCTTCGCTGAGCTTTCTCT 3120
      |||||
Db 361 AGATTAGCTCGCTCTGAGGAAGCGCCCTACAGCGGTTGCTTCGCTGAGCTTTCTCT 420

QY 3121 AGATGCTGCTGCTTCTCCTTCTCCTTCTCCTTCTTCTTCTTCTTCTTCTTCTTCT 3180
      |||||
Db 421 AGATGCTGCTGCTTCTCCTTCTCCTTCTCCTTCTTCTTCTTCTTCTTCTTCTTCT 480

QY 3181 GCGGGAGAGCCCAATATGAGCTTTGTTGGTAGCC 3215
      |||||
Db 481 GTGGGAGAGCCCAATATGAGCTTTGTTGGTAGCC 515

RESULT 15
LOCUS      BM117950/c
DEFINITION L0858E03-3 NIA Mouse Newborn Brain cDNA Library Mus musculus cDNA
ACCESSION  BM117950
VERSION     1
KEYWORDS   EST.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 403)
Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.
Construction of long-transcript enriched cDNA libraries from
submicrogram amounts of total RNAs by a universal PCR amplification
```

method  
Genome Res. 11 (9), 1553-1558 (2001)  
21429098  
11544199  
On Nov 26, 2001 this sequence version replaced gi:17080968.  
Contact: Dawood B. Dudekula  
Laboratory of Genetics  
National Institute on Aging/National Institutes of Health  
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
Email: cdna@lgsun.grc.nia.nih.gov  
Plate: L0858 row: E column: 03  
Seq primer: -21M13 Forward  
High quality sequence stop: 403  
POLYA=yes.

Location/Qualifiers  
1. .403  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="niaEST:L0858E03-3"  
/db\_xref="taxon:10090"  
/clone="L0858E03"  
/tissue\_type="Newborn Brain"  
/dev\_stage="Newborn"  
/lab\_host="DH10B"  
/clone\_lib="NIA Mouse Newborn Brain cDNA Library"  
/note="Organ: Brain; Vector: pSPORT1 (Invitrogen); Site 1:  
SalI; Site 2: NotI; Mouse cDNA project by the Laboratory  
of Genetics, National Institute on Aging (NIA), Intramural  
Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).  
Double-stranded cDNAs were synthesized with an Oligo(dT)  
primer [Invitrogen: 5'-  
pACTAGTCTAGATCGGAGCGGCGCCCTTTTCTTTTCTTTT-3'] from 48  
microgram of total RNA, treated with T4 DNA polymerase,  
and purified by ethanol-precipitation. The cDNAs were  
ligated to Lone-linker Li-Sal3 (Ref. Development  
127:1737-1749 (2000) [PMID:10725249]), purified by  
phenol/chloroform, and separated from free linkers by  
Centricon 100. Then, the cDNAs were digested with SalI and  
NotI enzymes, and cloned into SalI/NotI site of pSPORT1  
plasmid vector. The DH10B E. Coli host was transformed  
with ligation mixture by the standard chemical method. The  
average insert size is about 1.9 kb. The library was  
constructed by Yulan Piao (NIA)."

BASE COUNT 118 a 83 c 89 g 113 t

Query Match 10.5%; Score 361.6; DB 12; Length 403;  
Best Local Similarity 95.5%; Pred. No. 1.1e-36;  
Matches 384; Conservative 0; Mismatches 10; Indels 8; Gaps 1;

QY 3034 CCTAGCAGCTGTAGACCCAGAGCCAGGATTAGCTCGCTCTGAGGAGCGCCCTTACA 3093  
 |||||  
Db 403 CCTAGCAGCTGTAGACCCAGAGCCAGGATTAGCTCGCTCTGAGGAGCGCCCTTACA 344

QY 3094 GCGCGTGTCTCGTGGACTTTTCTAGATGCTGCTGCCATTACTCCAAAGTACTTC 3153  
 |||||  
Db 343 GCGCGTGTCTCGTGGACTTTTCTAGATGCTGCTGCCATTACTCCAAAGTACTTC 284

QY 3154 TATAAATCAAAACCTCTCTCCACAGCGGAGAGCCAAATATGAGCTTTGTTGGTAG 3213  
 |||||  
Db 283 TATAAATCAAAACCTCTCTCCACAGGTGGAGAGCCAAATATGAGCTTTGTTGGTAG 224

QY 3214 CCCGCTACCTTGGGGGCGCTT-----TCCACGAGCTTGAGGGGAAAGCCATGTATCT 3265  
 |||||  
Db 223 CCCGCTACCTTGGGGGCGCTT-----TCCAGGCGCCCGCCAGGCTTGAGGGGAAAGCCATGTATCT 164

QY 3266 GAAATATAGTATATCTTCTGTAATAGCTGAACAAACAAACCCGTTTTCGTAGAGGA 3325  
 |||||  
Db 163 GAAATATAGTATATCTTCTGTAATAGCTGAACAAACAAACCCGTTTTCGTAGAGGA 104

QY 3326 AAGCTAAATATCATTTTAAATATCTATGTTTAAATATCTATGTAACCTTTTTCATCTAT 3385  
 |||||

Db 103 AAGCTAAATATGATTTTAAATAATCTATGCTTTTAAATACTATGTAACCTTTTTCATCTAT 44

Qy 3386 TTAGTGATATATTTATGATGGAATAAACTTCTCTACTGTA 3427

Db 43 TTAGTGATATATTTATGATGGAATAAACTTCTCTACTGTA 2

Search completed: August 28, 2003, 05:42:11  
Job time : 8349.07 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 14:22:21 ; Search time 1028.85 Seconds  
(without alignments)  
9059.793 Million cell updates/sec

Title: US-09-919-408A-1  
Perfect score: 3453  
Sequence: 1 GCGCGCTGCTACCGCGC.....AAAAAAAAAAAAAAAAAAAA 3453

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349713017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Genesec\_19Jun03.\*

- 1: /SIDSI/gcgdata/genesec/genesecq-n-emb1/NA1980.DAT.\*
- 2: /SIDSI/gcgdata/genesec/genesecq-n-emb1/NA1981.DAT.\*
- 3: /SIDSI/gcgdata/genesec/genesecq-n-emb1/NA1982.DAT.\*
- 4: /SIDSI/gcgdata/genesec/genesecq-n-emb1/NA1983.DAT.\*
- 5: /SIDSI/gcgdata/genesec/genesecq-n-emb1/NA1984.DAT.\*
- 6: /SIDSI/gcgdata/genesec/genesecq-n-emb1/NA1985.DAT.\*
- 7: /SIDSI/gcgdata/genesec/genesecq-n-emb1/NA1986.DAT.\*
- 8: /SIDSI/gcgdata/genesec/genesecq-n-emb1/NA1987.DAT.\*
- 9: /SIDSI/gcgdata/genesec/genesecq-n-emb1/NA1988.DAT.\*
- 10: /SIDSI/gcgdata/genesec/genesecq-n-emb1/NA1989.DAT.\*
- 11: /SIDSI/gcgdata/genesec/genesecq-n-emb1/NA1990.DAT.\*
- 12: /SIDSI/gcgdata/genesec/genesecq-n-emb1/NA1991.DAT.\*
- 13: /SIDSI/gcgdata/genesec/genesecq-n-emb1/NA1992.DAT.\*
- 14: /SIDSI/gcgdata/genesec/genesecq-n-emb1/NA1993.DAT.\*
- 15: /SIDSI/gcgdata/genesec/genesecq-n-emb1/NA1994.DAT.\*
- 16: /SIDSI/gcgdata/genesec/genesecq-n-emb1/NA1995.DAT.\*
- 17: /SIDSI/gcgdata/genesec/genesecq-n-emb1/NA1996.DAT.\*
- 18: /SIDSI/gcgdata/genesec/genesecq-n-emb1/NA1997.DAT.\*
- 19: /SIDSI/gcgdata/genesec/genesecq-n-emb1/NA1998.DAT.\*
- 20: /SIDSI/gcgdata/genesec/genesecq-n-emb1/NA1999.DAT.\*
- 21: /SIDSI/gcgdata/genesec/genesecq-n-emb1/NA2000.DAT.\*
- 22: /SIDSI/gcgdata/genesec/genesecq-n-emb1/NA2001A.DAT.\*
- 23: /SIDSI/gcgdata/genesec/genesecq-n-emb1/NA2001B.DAT.\*
- 24: /SIDSI/gcgdata/genesec/genesecq-n-emb1/NA2002.DAT.\*
- 25: /SIDSI/gcgdata/genesec/genesecq-n-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3453	100.0	3453	16	Flk2 receptor prot
2	3453	100.0	3453	17	Human foetal liver
3	3453	100.0	3453	18	Murine flk-2 recep
4	3453	100.0	3453	20	Murine flk-2 cDNA.
5	3453	100.0	3453	24	Mouse cDNA encodin
6	3451.4	100.0	3453	14	Murine flk-2 cDNA.
7	3451.4	100.0	3453	14	Murine flk-2 codin
8	3451.4	100.0	3453	16	Mouse flk-2 cDNA.

9	3448.2	99.9	3453	13	AAQ29954	Murine flk-2 cDNA.
10	3446.6	99.8	3453	14	AAQ40914	Murine flk-2 cDNA.
11	3357	97.2	3440	24	AD42483	Murine receptor pr
12	3342.4	96.8	3521	16	AAQ0801	Flk2/filt3 tyrosine
13	2227.2	64.5	3501	16	AAQ81013	Flk2 receptor prot
14	2227.2	64.5	3501	16	AAQ79069	Human flk-2 cDNA.
15	2227.2	64.5	3501	17	AAT38734	Murine foetal live
16	2227.2	64.5	3501	18	AAT72117	Human flk-2 recept
17	2227.2	64.5	3501	20	AAQ77515	Human flk-2 cDNA.
18	2227.2	64.5	3501	24	ABSS5045	Human cDNA encodin
19	2225.6	64.5	3501	14	AAQ3503	Human flk-2 cDNA.
20	2225.6	64.5	3501	14	AAQ40915	Human flk-2 cDNA.
21	2224	64.4	3501	14	AAQ35250	Human flk-2 cDNA.
22	2205.4	63.9	3489	24	AD42484	Human receptor pro
23	2204.8	63.9	3476	16	AAQ81536	Human receptor pr
24	2198.2	63.7	3475	16	AAQ0802	Human flk2/filt3 ty
25	2085	60.4	3120	14	AAQ49756	Protein tyrosine-k
26	2081.8	60.3	3120	16	AAT03096	Human receptor typ
27	2024.6	58.6	2949	19	AAV39041	Human receptor typ
28	2018.8	58.5	2958	19	AAV39042	Human receptor typ
29	2017.2	58.4	2958	19	AAV39040	Human receptor typ
30	1994.8	57.8	2982	19	AAV39039	Human receptor typ
31	1984.2	57.5	2978	19	AAV39038	Flk-2ws gene. Mus
32	1199.6	34.7	1894	15	AAQ54036	Sequence pMON32390
33	1119.6	32.4	2247	19	AAV55294	DNA encoding novel
34	357.4	10.4	3992	23	AAQ79666	Breast carcinoma r
35	357.4	10.4	3992	24	ABV94301	Human macrophage c
36	357.4	10.4	3992	24	ABK48105	Human gene express
37	357.4	10.4	4008	24	ABK235433	Bovine c-Kit bk-1
38	332.6	9.6	3069	21	AAQ44718	Hampshire porcine
39	331.4	9.6	2919	20	AAV80687	Human c-kit oncoge
40	329	9.5	5084	19	AAV20443	Human Kit/stem cel
41	329	9.5	5084	22	AAQ13426	Breast cancer rela
42	329	9.5	5084	24	ABL64113	Ovary cancer relat
43	329	9.5	5084	24	ABL68085	DNA encoding novel
44	325.4	9.4	6390	23	AAQ79665	DNA encoding novel
45	325.4	9.4	6390	23	AAQ84936	

ALIGNMENTS

RESULT 1	
AAQ81012	
ID	AAQ81012 standard; cDNA; 3453 BP.
XX	
AC	AAQ81012;
XX	
DT	25-MAR-2003 (updated)
DT	16-AUG-1995 (first entry)
XX	
DE	Flk2 receptor protein-tyrosine-kinase cDNA.
XX	
KW	Mouse Flk2; receptor protein-tyrosine-kinase; primitive
KW	hematopoietic cell; fetal liver kinase; ds.
XX	
OS	Mus musculus.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	31..3009
FT	/*tag= a
FT	/product= Flk2 receptor protein-tyrosine-kinase
FT	31..111
FT	/*tag= b
FT	112..3006
FT	/*tag= c

XX	WO9500554-A2.
XX	05-JAN-1995.
XX	17-JUN-1994; 94WO-US06944.
XX	





Db 1681 CTCGTCTCCCTTCATTGTTCTCATTTGTTGATCTGCCACAATAACAAAAGCAA 1740  
Qy 1741 TTTAGGTACGAGAGTACGCTGAGATGATCCAGGTGACGTGGCCCCCTCGATACCGAGTAC 1800  
Db 1741 TTTAGGTACGAGAGTACGCTGAGATGATCCAGGTGACGTGGCCCCCTCGATACCGAGTAC 1800  
Qy 1801 TTTCTAGTTGACTTCAGGAGCTATGAATATGACCTTAAGTGGAGTCCCGAGAGAGAAC 1860  
Db 1801 TTTCTAGTTGACTTCAGGAGCTATGAATATGACCTTAAGTGGAGTCCCGAGAGAGAAC 1860  
Qy 1861 TTAGAGTTTGGAGAGTCTGGGGTCTGGCGCTTTCCGGAGGGTGAATGAACCCACCGGCC 1920  
Db 1861 TTAGAGTTTGGAGAGTCTGGGGTCTGGCGCTTTCCGGAGGGTGAATGAACCCACCGGCC 1920  
Qy 1921 TATGCACTTAGTAAACCGGAGTCTCAATTCAGGTGGCGTGAAGATGCTAAAGAGAAA 1980  
Db 1921 TATGCACTTAGTAAACCGGAGTCTCAATTCAGGTGGCGTGAAGATGCTAAAGAGAAA 1980  
Qy 1981 GCTGACAGTGTGAAGAGAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2040  
Db 1981 GCTGACAGTGTGAAGAGAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2040  
Qy 2041 CACCATGACCAATCGTGAATCTGCTGGGGGATGCACACTGTCCAGGGCCAGTGTACTTTG 2100  
Db 2041 CACCATGACCAATCGTGAATCTGCTGGGGGATGCACACTGTCCAGGGCCAGTGTACTTTG 2100  
Qy 2101 ATTTTGTGAATTTGTTGCTATGTTGATGCTCTCACTACCTAAGAGTAAAGAGAGAAAG 2160  
Db 2101 ATTTTGTGAATTTGTTGCTATGTTGATGCTCTCACTACCTAAGAGTAAAGAGAGAAAG 2160  
Qy 2161 TTTCACAGGACATGACAGAGATTTTAAAGAACATAATTTTCAGTTCTTACCCCTACTTTC 2220  
Db 2161 TTTCACAGGACATGACAGAGATTTTAAAGAACATAATTTTCAGTTCTTACCCCTACTTTC 2220  
Qy 2221 CAGGACATTCAAATTCAGCATGCTGCTGGGGGATGCACACTGTCCAGGGCCAGTGTACTTTG 2280  
Db 2221 CAGGACATTCAAATTCAGCATGCTGCTGGGGGATGCACACTGTCCAGGGCCAGTGTACTTTG 2280  
Qy 2281 GATCAGCTCTCAGGTTCAATGGGAATTCATTTCTGAAGATGAGATGGAATGAA 2340  
Db 2281 GATCAGCTCTCAGGTTCAATGGGAATTCATTTCTGAAGATGAGATGGAATGAA 2340  
Qy 2341 AACCCAGAGAGGCTGCGAGAGAGAGAGAGATTTGAACTGCTGAGAGTTGAGAC 2400  
Db 2341 AACCCAGAGAGGCTGCGAGAGAGAGAGAGATTTGAACTGCTGAGAGTTGAGAC 2400  
Qy 2401 CTCCTTTGCTTTGCTGCTACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAGTCTGT 2460  
Db 2401 CTCCTTTGCTTTGCTGCTACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAGTCTGT 2460  
Qy 2461 GTCCACAGAGACCTGCGAGCAGGAATGCTGCTGTCACCCACGGGAAGGTGGTGAAGATC 2520  
Db 2461 GTCCACAGAGACCTGCGAGCAGGAATGCTGCTGTCACCCACGGGAAGGTGGTGAAGATC 2520  
Qy 2521 TGTGACTTTGGACTGGCCCGAGACATCTCAGCGAGCTCCAGTACGTCGTCCAGGGCAAC 2580  
Db 2521 TGTGACTTTGGACTGGCCCGAGACATCTCAGCGAGCTCCAGTACGTCGTCCAGGGCAAC 2580  
Qy 2581 GCACGGCTGCGGTGAAGTGGATGGACCCCGAGAGCTTATTTGAAGGGATCTACACATC 2640  
Db 2581 GCACGGCTGCGGTGAAGTGGATGGACCCCGAGAGCTTATTTGAAGGGATCTACACATC 2640  
Qy 2641 AAGAGTGAAGTCTGCTGCTACGAGCATCTCTCTGGGAGATATTTTCACTGGGTGTAAC 2700  
Db 2641 AAGAGTGAAGTCTGCTGCTACGAGCATCTCTCTGGGAGATATTTTCACTGGGTGTAAC 2700  
Qy 2701 CCTTACCTGGCATTCCTGTCGAGCGTAACTTTCTATAACTGATTCAGAGTGAATTTAA 2760  
Db 2701 CCTTACCTGGCATTCCTGTCGAGCGTAACTTTCTATAACTGATTCAGAGTGAATTTAA 2760  
Qy 2761 ATGGAGCAGCCATTCATGCCCAGAGAGGATATACCTTTGTAATGCAATCTCTGCTGGCT 2820  
Db 2761 ATGGAGCAGCCATTCATGCCCAGAGAGGATATACCTTTGTAATGCAATCTCTGCTGGCT 2820

Qy 2821 TTTGACTCAAGGAAGCGGCCATCTTCCCAACCTGACTTCTATTTTTAGGATGTGAGCTG 2880  
Db 2821 TTTGACTCAAGGAAGCGGCCATCTTCCCAACCTGACTTCTATTTTTAGGATGTGAGCTG 2880  
Qy 2881 GCAGAGGAGAGAGAGATGATATCAGAACATCTCATCTCAACAAACAGCGGGCCCT 2940  
Db 2881 GCAGAGGAGAGAGAGATGATATCAGAACATCTCATCTCAACAAACAGCGGGCCCT 2940  
Qy 2941 CAGCAGAGAGCGGGCTCAGAGCCCGAGTGGCCACAGCGCGGATTCACAGAGAA 3000  
Db 2941 CAGCAGAGAGCGGGCTCAGAGCCCGAGTGGCCACAGCGCGGATTCACAGAGAA 3000  
Qy 3001 AGAGTGTAGCAGAGAGCGCTTGGACCCCGCCACCTAGCAGGCTGTAGACCGCAGAGCCA 3060  
Db 3001 AGAGTGTAGCAGAGAGCGCTTGGACCCCGCCACCTAGCAGGCTGTAGACCGCAGAGCCA 3060  
Qy 3061 AGATTAGCTCTGCTCTGAGGAAGCGCCCTACAGCGGTTGCTTCGCTGGAATTTTCTCT 3120  
Db 3061 AGATTAGCTCTGCTCTGAGGAAGCGCCCTACAGCGGTTGCTTCGCTGGAATTTTCTCT 3120  
Qy 3121 AGATGCTGTCTGCCATTTACTCCAAAGTGAATTTCTATAAATCAAACTCTCTCGCACAG 3180  
Db 3121 AGATGCTGTCTGCCATTTACTCCAAAGTGAATTTCTATAAATCAAACTCTCTCGCACAG 3180  
Qy 3181 GCGGGAGGAGCAATTAATGAGACTTGTGTGAGCCCGCTACCTCGGGGCTTTCCACG 3240  
Db 3181 GCGGGAGGAGCAATTAATGAGACTTGTGTGAGCCCGCTACCTCGGGGCTTTCCACG 3240  
Qy 3241 ACCTTGAGGGGAAAGCCATGATCTGAAATATAGTATATTCTTGTAAATACGTGAAACAA 3300  
Db 3241 ACCTTGAGGGGAAAGCCATGATCTGAAATATAGTATATTCTTGTAAATACGTGAAACAA 3300  
Qy 3301 ACCAAACCCGTTTTTGTCTAAGGAAAGCTAAATATGATTTTAAAAATCTATGTTTAA 3360  
Db 3301 ACCAAACCCGTTTTTGTCTAAGGAAAGCTAAATATGATTTTAAAAATCTATGTTTAA 3360  
Qy 3361 AATACATGTAATTTTCTATCTATTTAGTATATTTTATGATGGAATTAACCTTTC 3420  
Db 3361 AATACATGTAATTTTCTATCTATTTAGTATATTTTATGATGGAATTAACCTTTC 3420  
Qy 3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3453  
Db 3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3453

RESULT 2  
AAT38733  
ID AAT38733 standard; cDNA; 3453 BP.  
XX AAT38733;  
XX AC  
XX 25-MAR-2003 (updated)  
DT 11-DEC-1996 (first entry)  
XX Human foetal liver kinase 2 cDNA.  
DE Human; foetal liver kinase 2; flk-2; protein tyrosine kinase;  
KW monoclonal; antibody; extracellular domain; receptor assay;  
KW haematopoietic stem cell; ligand; stimulation; proliferation;  
KW differentiation; treatment; anaemia; bone marrow damage;  
KW cancer chemotherapy; radiation; ds.  
XX Homo sapiens.  
OS  
XX  
XX Key Location/Qualifiers  
FH CDS 31..3009  
FT /\*tag= a  
FT sig\_peptide 31..111  
FT /\*tag= b  
FT mat\_peptide 112..3006  
FT /\*tag= c  
XX

US5548065-A.  
 20-AUG-1996.  
 31-OCT-1994; 94US-0252517.  
 19-NOV-1992; 92US-0977451.  
 02-APR-1991; 91US-0679666.  
 28-JUN-1991; 91US-0728913.  
 15-NOV-1991; 91US-0793065.  
 24-DEC-1991; 91US-0813593.  
 26-JUN-1992; 92US-0906397.  
 12-NOV-1992; 92US-0975049.  
 30-APR-1993; 93US-0055269.  
 31-OCT-1994; 94US-0252517.  
 (UYPR-) UNIV PRINCETON.  
 Lemischka IR,  
 WPI; 1996-392678/39.  
 P-PSDB; AAR97418.  
 Anti-fœtal liver kinase 2 (flk-2) antibodies - useful in assays,  
 for isolating haematopoietic stem cells expressing receptor and for  
 obtaining ligands  
 Disclosure; Columns 27-34; 50pp; English.  
 The present sequence encodes human foetal liver kinase 2 (flk-2), a  
 protein tyrosine kinase. Isolated antibodies, pref. monoclonal,  
 raised against the extracellular portion of flk-2 can be used to  
 assay for flk receptors on the surface of primitive haematopoietic  
 stem cells, and to isolate positive cells. The antibodies can also  
 be used as, or to obtain ligands, which stimulate the proliferation  
 and/or differentiation of stem cells. The ligands can be used, e.g.  
 for treating anaemia, or bone marrow damage resulting from cancer  
 chemotherapy, or radiation.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX  
 SQ Sequence 3453 BP; 947 A; 822 C; 875 G; 809 T; 0 other;  
 Query Match 100.0%; Score 3453; DB 17; Length 3453;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GCGGCTGGCTACCGCGCTCCGGAGCCATGGGCGTTGGCGCAGCGCAGCGCGG 60  
 DB 1 GCGGCTGGCTACCGCGCTCCGGAGCCATGGGCGTTGGCGCAGCGCAGCGCGG 60  
 QY 61 CGGCTGCTGCTGT 120  
 DB 61 CGGCTGCTGCTGT 120  
 QY 121 CTGCTGTGATCAAGT 180  
 DB 121 CTGCTGTGATCAAGT 180  
 QY 181 CCATCATCGTACCGAATGT 240  
 DB 181 CCATCATCGTACCGAATGT 240  
 QY 241 CAGAGTGAAGGACGGTATATGAAGCGGCCACCGTGGAGTGGCGGAGTCTGGGTCCATC 300  
 DB 241 CAGAGTGAAGGACGGTATATGAAGCGGCCACCGTGGAGTGGCGGAGTCTGGGTCCATC 300  
 QY 301 ACCCTGCAAGTGTGCTGCCACCCAGGGGACCTTCTGTGCTCTGTGGGTCTTTAAGCAC 360  
 DB 301 ACCCTGCAAGTGTGCTGCCACCCAGGGGACCTTCTGTGCTCTGTGGGTCTTTAAGCAC 360  
 QY 361 AGCTCCCTGGCTGCCAGCGGCACTTTGATTTTACAAACACAGAGGATCGTTCCATGGCC 420  
 DB 361 AGCTCCCTGGCTGCCAGCGGCACTTTGATTTTACAAACACAGAGGATCGTTCCATGGCC 420

QY 421 ATCTTGAACGTGACAGAGACCCAGGAGAGTAATACCTACTCCATATTTTCAGAGCGAACGC 480  
 DB 421 ATCTTGAACGTGACAGAGACCCAGGAGAGTAATACCTACTCCATATTTTCAGAGCGAACGC 480  
 QY 481 GCCAACTACACAGTACTGTTTACAGTGAATGTAAGAGATACACAGCTGTATGTGCTAAGG 540  
 DB 481 GCCAACTACACAGTACTGTTTACAGTGAATGTAAGAGATACACAGCTGTATGTGCTAAGG 540  
 QY 541 AGACCTTACTTTAGGAAGTGAAGAACCCAGGATGCACTGCTCTGCTATCTCCAGGGTGT 600  
 DB 541 AGACCTTACTTTAGGAAGTGAAGAACCCAGGATGCACTGCTCTGCTATCTCCAGGGTGT 600  
 QY 601 CCGAGGCCCACTGTGGAGTGGTGTCTGTGAGTCTCCACAGGAAAGCTGTAAAGAA 660  
 DB 601 CCGAGGCCCACTGTGGAGTGGTGTCTGTGAGTCTCCACAGGAAAGCTGTAAAGAA 660  
 QY 661 GGCCCTGCTGTGTGTCAGAAAGGAGGAAAGGTAATTCATGAGTGTTCGGAAACAGACATC 720  
 DB 661 GGCCCTGCTGTGTGTCAGAAAGGAGGAAAGGTAATTCATGAGTGTTCGGAAACAGACATC 720  
 QY 721 AGATGCTGTGTAGAAATGCACTGGGCGCGAATGCACTGAGTGTTCACCATATGATCTA 780  
 DB 721 AGATGCTGTGTAGAAATGCACTGGGCGCGAATGCACTGAGTGTTCACCATATGATCTA 780  
 QY 781 AACAGGCTCTCTAGAGCACAACCTCCAGTATTTCTGAAAGTGGGGGAAACCTTTGG 840  
 DB 781 AACAGGCTCTCTAGAGCACAACCTCCAGTATTTCTGAAAGTGGGGGAAACCTTTGG 840  
 QY 841 ATCAGGTTTAAAGCCATTCATGTGAACCATGATTTCCGGCTCACCTGGGAGCTGGAAGAC 900  
 DB 841 ATCAGGTTTAAAGCCATTCATGTGAACCATGATTTCCGGCTCACCTGGGAGCTGGAAGAC 900  
 QY 901 AAAGCCCTGAGGAGGAGGAGCTACTTTGAGATGATGATCTTCTCCACAAAACAGACCATG 960  
 DB 901 AAAGCCCTGAGGAGGAGGAGCTACTTTGAGATGATGATCTTCTCCACAAAACAGACCATG 960  
 QY 961 ATTCCGATTTCTTCTGGCTTTTGTCTTCTCGTGGAGGAGGAGCAGACCGGATATTACAC 1020  
 DB 961 ATTCCGATTTCTTCTGGCTTTTGTCTTCTCGTGGAGGAGGAGCAGACCGGATATTACAC 1020  
 QY 1021 TGCTTCTTCTCAAGACACCCAGCAGCTAGCTGAGCTGTGTGAGTCTCTAGAAAAGGGTTT 1080  
 DB 1021 TGCTTCTTCTCAAGACACCCAGCAGCTAGCTGAGCTGTGTGAGTCTCTAGAAAAGGGTTT 1080  
 QY 1081 ATAAACGCTACAGCTCGAAGAGATGTAATTTGACCGGTACGAAAGTTCTGCTTC 1140  
 DB 1081 ATAAACGCTACAGCTCGAAGAGATGTAATTTGACCGGTACGAAAGTTCTGCTTC 1140  
 QY 1141 TCAGTCAAGTTTAAAGCGTACCCAGCAATCCGATGACCGTGGATCTTCTCAAGCCTCA 1200  
 DB 1141 TCAGTCAAGTTTAAAGCGTACCCAGCAATCCGATGACCGTGGATCTTCTCAAGCCTCA 1200  
 QY 1201 TTTTCTTGTGAAACAGAGAGGCTGGAGGATGGGTACAGCATATCTAAATTTTGGGATCAT 1260  
 DB 1201 TTTTCTTGTGAAACAGAGAGGCTGGAGGATGGGTACAGCATATCTAAATTTTGGGATCAT 1260  
 QY 1261 AAGAAACAGCAGGAGATGATATTTCTATGCAAGAAATGATGACGCCAGTTTCAACAAA 1320  
 DB 1261 AAGAAACAGCAGGAGATGATATTTCTATGCAAGAAATGATGACGCCAGTTTCAACAAA 1320  
 QY 1321 ATGTTCCAGCTGAATATAAGAAAGAAACCTCAGTGTCTAGCAATGCTCAGCAGCCAG 1380  
 DB 1321 ATGTTCCAGCTGAATATAAGAAAGAAACCTCAGTGTCTAGCAATGCTCAGCAGCCAG 1380  
 QY 1381 GGTCTCTGTCTCTGATGGCTACCCGCTACCTCTTGGACCTTGGAGAGTGTTCGGAC 1440  
 DB 1381 GGTCTCTGTCTCTGATGGCTACCCGCTACCTCTTGGACCTTGGAGAGTGTTCGGAC 1440  
 QY 1441 AAATCTCCCAATGCAAGGAGAAATCCAGAGGAGTTTGGAAATAAAAGGCTAACAGA 1500  
 DB 1441 AAATCTCCCAATGCAAGGAGAAATCCAGAGGAGTTTGGAAATAAAAGGCTAACAGA 1500



haematopoietic hierarchy; extracellular domain; soluble form; ligand; proliferation; differentiation; mammalian; haematopoietic stem cell; macrocytic anaemia; aplastic anaemia; cancer; radiation; mouse; ss.  
Mus musculus.

Key Location/Qualifiers  
CDS 31..3009  
/\*tag= a  
/\*product= Murine\_flk-2  
sig\_peptide 31..111  
/\*tag= b  
mat\_peptide 112..3006  
/\*tag= c

US5621090-A.

15-APR-1997.

26-JUN-1992; 92US-0906397.

26-JUN-1992; 92US-0906397.

02-APR-1991; 91US-0679666.

28-JUN-1991; 91US-0728913.

15-NOV-1991; 91US-0793065.

24-DEC-1991; 91US-0813593.

(UYP- ) UNIV PRINCETON.

Lemischka IR;

WPI; 1997-235228/21.

P-PSDB; AAW19874.

Protein containing the extracellular domain of human flk-2 - used for identification of primitive haematopoietic cell proliferation and differentiation stimulatory ligands, e.g. for treating anaemia  
Disclosure; Fig 1A; 55pp; English.

This sequence encodes the murine fetal liver kinase 2 (flk2). flk-2 is a receptor protein tyrosine kinase (PTK) and is important in transducing putative self-renewal signals from the environment. flk-2 is expressed in fetal liver, spleen, fetal thymus, adult brain and adult bone marrow, and it is thought that flk-2 is expressed in the entire primitive portion of the haematopoietic hierarchy. The invention concerns a recombinant nucleic acid, preferably mRNA, which encodes a protein containing only the extracellular domain of human flk-2 and lacking the flk-2 intracellular catalytic domain. The resultant protein represents a soluble form of flk-2 which is used to isolate specific ligands for flk-2. These ligands can be used to stimulate proliferation and/or differentiation of mammalian haematopoietic stem cells, in vivo or in vitro, e.g. for treatment of macrocytic or aplastic anaemia or bone marrow damage caused by cancer treatment or radiation.  
(Updated on 25-MAR-2003 to correct PF field.)

Sequence 3453 BP; 947 A; 822 C; 875 G; 809 T; 0 other;

Query Match 100.0%; Score 3453; DB 18; Length 3453;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GCGGCTGGCTACCGCGGCTCCGGAGCCGATGCGGCGTGGCGGAGCGGACCGG 60  
1 GCGGCTGGCTACCGCGGCTCCGGAGCCGATGCGGCGTGGCGGAGCGGACCGG 60  
61 CGGCTGCTGCTGCTGTTGTTGTCAGTAATGATTTCTGAGACCGTTTCAAAACCAAGAC 120  
61 CGGCTGCTGCTGCTGTTGTTGTTGTCAGTAATGATTTCTGAGACCGTTTCAAAACCAAGAC 120  
121 CTGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAAACAATGGCTCATCAGCGGGGAAG 180  
121 CTGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAAACAATGGCTCATCAGCGGGGAAG 180

QY 181 CCATCATCGTACCGAATGGTGGAGGATCCCGAAGACCTCCAGTGTACCCCGAGGCGC 240  
DB 181 CCATCATCGTACCGAATGGTGGAGGATCCCGAAGACCTCCAGTGTACCCCGAGGCGC 240  
QY 241 CAGAGTGAAGGACGGGTATATGAAGGGCCACCGTGGAGGTGGCCAGTCTGGGTCCATC 300  
DB 241 CAGAGTGAAGGACGGGTATATGAAGGGCCACCGTGGAGGTGGCCAGTCTGGGTCCATC 300  
QY 301 ACCCTCAAGTGCAGCTCGCCACCCAGGGGACCTTCTGCTCTGGGTCTTTAAAGCAC 360  
DB 301 ACCCTCAAGTGCAGCTCGCCACCCAGGGGACCTTCTGCTCTGGGTCTTTAAAGCAC 360  
QY 361 AGCTCCCTGGGCTGCCAGCGGCACTTTGATTTTACAAAACAGAGGAATCGTTTCAATGGCC 420  
DB 361 AGCTCCCTGGGCTGCCAGCGGCACTTTGATTTTACAAAACAGAGGAATCGTTTCAATGGCC 420  
QY 421 ATCTTGAACGTGACAGAGACCCAGGAGGAGATACCTACTCCATATTCAGAGCGAACGC 480  
DB 421 ATCTTGAACGTGACAGAGACCCAGGAGGAGATACCTACTCCATATTCAGAGCGAACGC 480  
QY 481 GCCAACTACACAGTACTGTTTCAAGTGAATGTAAGAGATACACAGCTGTATGTCTAAGG 540  
DB 481 GCCAACTACACAGTACTGTTTCAAGTGAATGTAAGAGATACACAGCTGTATGTCTAAGG 540  
QY 541 AGACCTTACTTTAGGAAGATGGAACCCAGGATGCACTGCTCTGCACTCCCGAGGGTGT 600  
DB 541 AGACCTTACTTTAGGAAGATGGAACCCAGGATGCACTGCTCTGCACTCCCGAGGGTGT 600  
QY 601 CCGAGGCCACTGTGGAGTGGTGTCTCTGCAGCTCCACAGGGAAAGCTGTAAAGAGAA 660  
DB 601 CCGAGGCCACTGTGGAGTGGTGTCTCTGCAGCTCCACAGGGAAAGCTGTAAAGAGAA 660  
QY 661 GGCCCTGCTGTTTTCAGAAAAGGAGGAAAGGTAATTCTATAGTGTGTCGAAACAGACATC 720  
DB 661 GGCCCTGCTGTTTTCAGAAAAGGAGGAAAGGTAATTCTATAGTGTGTCGAAACAGACATC 720  
QY 721 AGATGCTGTCTAGAAATGCACTGGCGCGGAGTGCACCAAGCTGTTCACATAGATCTA 780  
DB 721 AGATGCTGTCTAGAAATGCACTGGCGCGGAGTGCACCAAGCTGTTCACATAGATCTA 780  
QY 781 AACAGGCTCTCAGAGACACACTGCCCCAGTATTTCTGAAAGTGGGGGAAACCTTTGG 840  
DB 781 AACAGGCTCTCAGAGACACACTGCCCCAGTATTTCTGAAAGTGGGGGAAACCTTTGG 840  
QY 841 ATCAGTGTAAAGCCATTCATGTGAACCATGATGCGGCTCACTGGGAGCTGGAAGAC 900  
DB 841 ATCAGTGTAAAGCCATTCATGTGAACCATGATGCGGCTCACTGGGAGCTGGAAGAC 900  
QY 901 AAAGCCCTGGAGGAGGCGAGTACTTTGAGATGATGATCTACTCAAAACAGGACCATG 960  
DB 901 AAAGCCCTGGAGGAGGCGAGTACTTTGAGATGATGATCTACTCAAAACAGGACCATG 960  
QY 961 ATTGGGATTTCTTTGGCTTTTGTCTTCCGTGGGAAGGAGACGACCGGATATATACCC 1020  
DB 961 ATTGGGATTTCTTTGGCTTTTGTCTTCCGTGGGAAGGAGACGACCGGATATATACCC 1020  
QY 1021 TGCTCTTCTTCAAGACACCCAGCAGTCAAGCTGTGGTGGTACCCTCTAGAAAAAGGGTTT 1080  
DB 1021 TGCTCTTCTTCAAGACACCCAGCAGTCAAGCTGTGGTGGTACCCTCTAGAAAAAGGGTTT 1080  
QY 1081 ATAAACGCTACAGCTCGGAAGAGATGAAATTTGAACCGTACGAAAGTTCCTGCTTC 1140  
DB 1081 ATAAACGCTACAGCTCGGAAGAGATGAAATTTGAACCGTACGAAAGTTCCTGCTTC 1140  
QY 1141 TCAGTCAGGTTTAAAGCGTACCCAGCAATCCGATGACCGTGGATCTTCTCTCAGCCCTCA 1200  
DB 1141 TCAGTCAGGTTTAAAGCGTACCCAGCAATCCGATGACCGTGGATCTTCTCTCAGCCCTCA 1200  
QY 1201 TTTCTTGTGAACAGAGAGCCCTGGAGGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260  
DB 1201 TTTCTTGTGAACAGAGAGCCCTGGAGGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260

QY 1261 AAGAACAGCCAGGAGTAGTACATATTTCTATGCAGAAAATGATGACGCCAGTTCAACAAA 1320  
DB 1261 AAGAAACAGCCAGGAGTAGTACATATTTCTATGCAGAAAATGATGACGCCAGTTCAACAAA 1320  
QY 1321 ATGTTCAAGCTGAATATAAGAAAGAAACCTCAAGTGTCTAGCAAAATGCTCAGCCAGCCAG 1380  
DB 1321 ATGTTCAAGCTGAATATAAGAAAGAAACCTCAAGTGTCTAGCAAAATGCTCAGCCAGCCAG 1380  
QY 1381 GCCTCTGTTTCTCTGATGCTACCCGCTACCCCTCTTGGACCTTGGAAAGAGTGTTCGGAC 1440  
DB 1381 GCCTCTGTTTCTCTGATGCTACCCGCTACCCCTCTTGGACCTTGGAAAGAGTGTTCGGAC 1440  
QY 1441 AATCTCCCAATTTGACGAGGAATCCAGAGGAGTTTGGATATAAAGGCTTAACAGA 1500  
DB 1441 AATCTCCCAATTTGACGAGGAATCCAGAGGAGTTTGGATATAAAGGCTTAACAGA 1500  
QY 1501 AAAGTGTTCGCGCAGTGGGTGTGAGCAGTACTCTAAATATGAGTAGGCGCGGAAGAGG 1560  
DB 1501 AAAGTGTTCGCGCAGTGGGTGTGAGCAGTACTCTAAATATGAGTAGGCGCGGAAGAGG 1560  
QY 1561 CTTCTGCTCAAAATGCTGTGCTGCTCAAAATTTCTATGGGCAAGTCTTTCGAAACCAATCTTTTA 1620  
DB 1561 CTTCTGCTCAAAATGCTGTGCTGCTCAAAATTTCTATGGGCAAGTCTTTCGAAACCAATCTTTTA 1620  
QY 1621 AACTCAGGAGCCCTTCCCTTTTCATCCAGAAACCAATCTTCTATGCGAACAATTTGGG 1680  
DB 1621 AACTCAGGAGCCCTTCCCTTTTCATCCAGAAACCAATCTTCTATGCGAACAATTTGGG 1680  
QY 1681 CTCTGCTCCCTTCAATTTGTTTCTCAATTTGTTGATCTGCGCAAAATACAAAGCAA 1740  
DB 1681 CTCTGCTCCCTTCAATTTGTTTCTCAATTTGTTGATCTGCGCAAAATACAAAGCAA 1740  
QY 1741 TTTAGGTACGAGAGTCAGCTGCAGATGATCCAGGTGACTGCGCCCTCGGATTAAGAGTAC 1800  
DB 1741 TTTAGGTACGAGAGTCAGCTGCAGATGATCCAGGTGACTGCGCCCTCGGATTAAGAGTAC 1800  
QY 1801 TTCTAGCTTGAATAGTAAATATGACCTTAAGTGGGAGTTCGCGAGAGAAAC 1860  
DB 1801 TTCTAGCTTGAATAGTAAATATGACCTTAAGTGGGAGTTCGCGAGAGAAAC 1860  
QY 1861 TTAGAGTTTGGGAGGTCCTGGGTCTGGCGCTTTCGGGAGGGTGAACGCCACCGCC 1920  
DB 1861 TTAGAGTTTGGGAGGTCCTGGGTCTGGCGCTTTCGGGAGGGTGAACGCCACCGCC 1920  
QY 1921 TATGGCATTAAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGTCTAAAGAGAAA 1980  
DB 1921 TATGGCATTAAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGTCTAAAGAGAAA 1980  
QY 1981 GCTGACAGCTGTGAAAAGAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA 2040  
DB 1981 GCTGACAGCTGTGAAAAGAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA 2040  
QY 2041 CACCATGACAAATCGTGAATCTCTGGGGCATGACATGTCAGGGCCAGTGTACTTG 2100  
DB 2041 CACCATGACAAATCGTGAATCTCTGGGGCATGACATGTCAGGGCCAGTGTACTTG 2100  
QY 2101 ATTTTGAATATTTGCTATGTTGACTCTCTCAACTACCTTAAGAGTAAAGAGAGAA 2160  
DB 2101 ATTTTGAATATTTGCTATGTTGACTCTCTCAACTACCTTAAGAGTAAAGAGAGAA 2160  
QY 2161 TTTTACAGGACATGGAACAGAGATTTTAAAGGAAATATTTTCACTTCTTACCTTCTTTC 2220  
DB 2161 TTTTACAGGACATGGAACAGAGATTTTAAAGGAAATATTTTCACTTCTTACCTTCTTTC 2220  
QY 2221 CAGGCAATTTCAAAATTCAGGATTCGCTGGTTCAGAGAGTTCAGTTCACCCGCTTTC 2280  
DB 2221 CAGGCAATTTCAAAATTCAGGATTCGCTGGTTCAGAGAGTTCAGTTCACCCGCTTTC 2280  
QY 2281 GATCAGCTCTCAGGTTTCAATGGGAAATTCATTTCAATTCAGAGATGAGATTTGATATGAA 2340  
DB 2281 GATCAGCTCTCAGGTTTCAATGGGAAATTCATTTCAATTCAGAGATGAGATTTGATATGAA 2340  
QY 2341 AACAGAGAGGCTGGCAGAGAGAGAGAGAGATTTGAACGTGCTGACGTTTGAAGAC 2400

DB 2341 AACAGAGAGGCTGGCAGAGAGAGAGAGAGATTTGAACGTGCTGACGTTTGAAGAC 2400  
QY 2401 CTCCTTTGCTTTCCGTAACCAAGTGGCCAAAGGCAATGGAATTCCTGGAGTTCAAGTCGTGT 2460  
DB 2401 CTCCTTTGCTTTCCGTAACCAAGTGGCCAAAGGCAATGGAATTCCTGGAGTTCAAGTCGTGT 2460  
QY 2461 GTCCACAGAGACCTGGCAGCCAGGAATGTTGGTCAACCCACGGAAGGTGGTCAAGATC 2520  
DB 2461 GTCCACAGAGACCTGGCAGCCAGGAATGTTGGTCAACCCACGGAAGGTGGTCAAGATC 2520  
QY 2521 TGTGACTTTTGGACTGGCCCGAGACATCTCAGCGACCTCCAGCTACGTCTGTCAGGGCAAC 2580  
DB 2521 TGTGACTTTTGGACTGGCCCGAGACATCTCAGCGACCTCCAGCTACGTCTGTCAGGGCAAC 2580  
QY 2581 GCACGGCTGCGGTGAAGTGGATGGCACCCGAGAGCTTTAATTGAAGGATCTACACAATC 2640  
DB 2581 GCACGGCTGCGGTGAAGTGGATGGCACCCGAGAGCTTTAATTGAAGGATCTACACAATC 2640  
QY 2641 AAGAGTGACGTCTGGTCTCTACGGCATCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700  
DB 2641 AAGAGTGACGTCTGGTCTCTACGGCATCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700  
QY 2701 CCTTACCTGGCATCTCTGTCGACGCTAACTTCTATAAACTGATTCAGAGTGAATTTAAA 2760  
DB 2701 CCTTACCTGGCATCTCTGTCGACGCTAACTTCTATAAACTGATTCAGAGTGAATTTAAA 2760  
QY 2761 ATGGAGCAGCCATTTCTATGTCACAGAGAGGATATACCTTGTAAATGCAATCTCTGCTGGCT 2820  
DB 2761 ATGGAGCAGCCATTTCTATGTCACAGAGAGGATATACCTTGTAAATGCAATCTCTGCTGGCT 2820  
QY 2821 TTTGACTCAAGGAGCGGCCATCTTCCCAACCTGACTTCAATTTTAGGATGTGAGCTG 2880  
DB 2821 TTTGACTCAAGGAGCGGCCATCTTCCCAACCTGACTTCAATTTTAGGATGTGAGCTG 2880  
QY 2881 GCAGAGCAGAGAGAGATGATCAGAAACATCCATCCATCTACCAAAACAGCGGCCCT 2940  
DB 2881 GCAGAGCAGAGAGAGATGATCAGAAACATCCATCCATCTACCAAAACAGCGGCCCT 2940  
QY 2941 CAGCAGAGAGCGGGCTCAGAGCCAGTCCGACAGCGCCAGCGTGAAGATTCACAGAGAA 3000  
DB 2941 CAGCAGAGAGCGGGCTCAGAGCCAGTCCGACAGCGCCAGCGTGAAGATTCACAGAGAA 3000  
QY 3001 AGAAGTTAGCGAGAGCGCTTGGACCCGCAACCTAGCAGGCTGTAGACCGCAGAGCA 3060  
DB 3001 AGAAGTTAGCGAGAGCGCTTGGACCCGCAACCTAGCAGGCTGTAGACCGCAGAGCA 3060  
QY 3061 AGATTAGCTCGCTCTGAGAGAGCGCTTACAGCGGTTGCTTCTGCTGGACTTTTCTCT 3120  
DB 3061 AGATTAGCTCGCTCTGAGAGAGCGCTTACAGCGGTTGCTTCTGCTGGACTTTTCTCT 3120  
QY 3121 AGATGCTGTCTGCCATTTACTTCAAAAGTGAATTTCTATAAAATCAAACTCTCTGACAG 3180  
DB 3121 AGATGCTGTCTGCCATTTACTTCAAAAGTGAATTTCTATAAAATCAAACTCTCTGACAG 3180  
QY 3181 GCGGGAGAGCAATAATAGAGACTTGTGGTGGAGCGCCCTACCCCTGGGGCTTTTCCACG 3240  
DB 3181 GCGGGAGAGCAATAATAGAGACTTGTGGTGGAGCGCCCTACCCCTGGGGCTTTTCCACG 3240  
QY 3241 AGCTTGAAGGGAAGCCATGATCTGAATATATAGTATATTTCTGTAAATACGTGAACAA 3300  
DB 3241 AGCTTGAAGGGAAGCCATGATCTGAATATATAGTATATTTCTGTAAATACGTGAACAA 3300  
QY 3301 ACCAAACCCGTTTTTCTAAGGGAAGCTTAAATATGATATTTTAAATCTATGTTTTAA 3360  
DB 3301 ACCAAACCCGTTTTTCTAAGGGAAGCTTAAATATGATATTTTAAATCTATGTTTTAA 3360  
QY 3361 AATACTATGTAATCTTTTCTATTTAGTGAATATTTTATGATGGAATAAATCTTTC 3420  
DB 3361 AATACTATGTAATCTTTTCTATTTAGTGAATATTTTATGATGGAATAAATCTTTC 3420  
QY 3421 TACTGTAAAAAAGGAGAGAGAGAGAGATTTGAACGTGCTGACGTTTGAAGAC 3453

Dd     3421 TACTGTAAAAA.....AAAAAAAAAAAAAAA 3453

RESULT 4  
AAAX77514

ID   AAAX77514 standard; cdNA; 3453 BP.

XX                 AC AAAX77514;

XX                 DT      05-AUG-1999   (first entry)

XX                 DE   Murine flk-2 cDNA.

XX                 KW   Murine; flk-2; flk-1; cell isolation; fetal liver kinase; receptor;  
monoclonal; polyclonal; antibody; tyrosine kinase; ds.

XX                 OS   Mus sp.

FH                  Key                   Location/Qualifiers  
FT                  CDS                   31..3009  
FT                   /\*tag= a  
FT                   /product= "flk-2"

XX                   US5912133-A.

XX                   PD   15-JUN-1999.

XX                   PF   10-FEB-1998;   98US-0021324.

PR   19-NOV-1992;   92US-0977451.  
PR   02-APR-1991;   91US-0679666.  
PR   28-JUN-1991;   91US-0728913.  
PR   15-NOV-1991;   91US-0793065.  
PR   24-DEC-1991;   91US-0813593.  
PR   26-JUN-1992;   92US-0906397.  
PR   12-NOV-1992;   92US-0975049.  
PR   30-APR-1993;   93US-0055269.  
PR   31-OCT-1994;   94US-0252498.  
PR   15-FEB-1996;   96US-0601891.

XX                   PA   (UYPR-) UNIV PRINCETON.

PI                   Lemischka IR;

DR   WPI; 1999-357194/30.  
DR   P-PADB; AAY08616.

XX                   PT   Isolating hematopoietic cells expressing fetal liver kinase 1  
PT                   receptors

PS   Disclosure; Fig la; 59pp; English.

CC                   This invention describes a novel method of isolating cells expressing  
fetal liver kinase 1 (flk-1) receptors on their surface and comprises  
binding the cells to a polyclonal or monoclonal antibody specific to  
the flk-1 receptor and isolating the cells that have bound to the  
antibody. The method can be used to isolate hematopoietic stem cells in  
any mammal but preferably a rat, mouse, rabbit or human. The proteins of  
the invention belong to the receptor protein family. This sequence  
encodes the murine flk-2 protein which is used in the method of the  
invention.

SQ   Sequence 3453 BP; 947 A; 822 C; 875 G; 809 T; 0 other;

Query Match                   100.0%; Score 3453; DB 20; Length 3453;  
Best Local Similarity       100.0%; Pred. No. 0;  
Matches 3453; Conservative   0; Mismatches   0; Indels   0; Gaps   0

Oy   1 GCGCCTGGTACC CGCGGCTTC CGAGGCCATG CGGGCGTT GGCGC AGCGAGCG ACCGG 60  
| | | | |  
Db   1 GCGCCTGGTACC CGCGGCTTC CGAGGCCATG CGGGCGTT GGCGC AGCGAGCG ACCGG 60  
| | | | |

Oy   61 CGGCTGCTGCT GTTGTTG TTGTC TAGTA ATCTTAGAC CGGTTACA AACCAAG AC 120  
| | | | |



Db 1141 TCAGTCAGGTTTAAAGCGTACCACGAATCCGATGCACGTGGATCTTCTCTCAGCCCTCA 1200  
 Qy 1201 TTTCTCTGTGAA CAGAGAGCCCTGGAGATGGGTACAGCATATCTAAATTTTTCGGATCAT 1260  
 Db 1201 TTTCTCTGTGAA CAGAGAGCCCTGGAGATGGGTACAGCATATCTAAATTTTTCGGATCAT 1260  
 Qy 1261 AAGAACAGCCAGGAGTACATATTTCTATGCAAAATGATGACGCCCGAGTTTACACAAA 1320  
 Db 1261 AAGAACAGCCAGGAGTACATATTTCTATGCAAAATGATGACGCCCGAGTTTACACAAA 1320  
 Qy 1321 ATGTTCACGCTGAATATAGAAAGAACTCAAGTGTCTAGCAATGCTCAGCCAGCCAG 1380  
 Db 1321 ATGTTCACGCTGAATATAGAAAGAACTCAAGTGTCTAGCAATGCTCAGCCAGCCAG 1380  
 Qy 1381 GGGTCTGTCTCTGATGGCTACCCGCTACCTCTTGGACCTTGGAAAGAGTGTTCGGAC 1440  
 Db 1381 GGGTCTGTCTCTGATGGCTACCCGCTACCTCTTGGACCTTGGAAAGAGTGTTCGGAC 1440  
 Qy 1441 AAATCTCCCAATTTGACGGAGGAATCCCAAGAGGAGTTTGGAAATAAAAGGCTTAACAGA 1500  
 Db 1441 AAATCTCCCAATTTGACGGAGGAATCCCAAGAGGAGTTTGGAAATAAAAGGCTTAACAGA 1500  
 Qy 1501 AAGAGTGTGGCCAGTGGGTGTCGACGAGTACTCTAAATATGATGAGGCGCGGGAAGGG 1560  
 Db 1501 AAGAGTGTGGCCAGTGGGTGTCGACGAGTACTCTAAATATGATGAGGCGCGGGAAGGG 1560  
 Qy 1561 CTTCTGCTCAAAATGCTGTGGTCAAAATTTCTATGGGCACGCTCTTGGCAAAACATCTTTTAA 1620  
 Db 1561 CTTCTGCTCAAAATGCTGTGGTCAAAATTTCTATGGGCACGCTCTTGGCAAAACATCTTTTAA 1620  
 Qy 1621 AACTCAGGAGCCCTTCCCTTCAATCCAAAGAACATCTCTCTATGCGACCATTTGGG 1680  
 Db 1621 AACTCAGGAGCCCTTCCCTTCAATCCAAAGAACATCTCTCTATGCGACCATTTGGG 1680  
 Qy 1681 CTCTGCTCCCTTCAATGTTGTTCTCATTTGTTGTTGATCTGCCACAAATACAAAAGCAA 1740  
 Db 1681 CTCTGCTCCCTTCAATGTTGTTCTCATTTGTTGTTGATCTGCCACAAATACAAAAGCAA 1740  
 Qy 1741 TTTAGGTACAGAGTCAAGTGCAGATGATCCAGGTGATCGGCCCTTGGATTAACAGTAC 1800  
 Db 1741 TTTAGGTACAGAGTCAAGTGCAGATGATCCAGGTGATCGGCCCTTGGATTAACAGTAC 1800  
 Qy 1801 TTCTACGTTTGAATTCAGGAGTATGAATATAGACCTTAAGTGGAGTTCCTCGAGAGAGAC 1860  
 Db 1801 TTCTACGTTTGAATTCAGGAGTATGAATATAGACCTTAAGTGGAGTTCCTCGAGAGAGAC 1860  
 Qy 1861 TTAGAGTTTGGGAAGTCTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1920  
 Db 1861 TTAGAGTTTGGGAAGTCTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1920  
 Qy 1921 TATGGCATAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAGAGAGAAA 1980  
 Db 1921 TATGGCATAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAGAGAGAAA 1980  
 Qy 1981 GCTGACAGCTGTGAAAAGAGTCTCTATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2040  
 Db 1981 GCTGACAGCTGTGAAAAGAGTCTCTATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2040  
 Qy 2041 CACCATGACAAATCTGTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 2100  
 Db 2041 CACCATGACAAATCTGTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 2100  
 Qy 2101 ATTTTGTGAATATGTTGCTATGGTGAACCTCTCACTACCTAAGAGTAAAGAGAGAG 2160  
 Db 2101 ATTTTGTGAATATGTTGCTATGGTGAACCTCTCACTACCTAAGAGTAAAGAGAGAG 2160  
 Qy 2161 TTTTACAGGACATGACAGAGATTTTAAAGGAATAAATTCAGTCTTCTTACCTACTTTC 2220  
 Db 2161 TTTTACAGGACATGACAGAGATTTTAAAGGAATAAATTCAGTCTTCTTACCTACTTTC 2220  
 Qy 2221 CAGGCACATTCAAATTCAGCATGCTGTTTACAGAGAGTTTACAGTACACCCGCCCTTG 2280  
 Db 2221 CAGGCACATTCAAATTCAGCATGCTGTTTACAGAGAGTTTACAGTACACCCGCCCTTG 2280

Qy 2281 GATCAGCTCTCAGGCTTCAATGGGAATTCAAATTCATTTCTGAAGATGAGATTGAATGAA 2340  
 Db 2281 GATCAGCTCTCAGGCTTCAATGGGAATTCAAATTCATTTCTGAAGATGAGATTGAATGAA 2340  
 Qy 2341 AACACAGAGAGCTGGCAGAAAGAGAGAGAGATTTGAAACGTGCTGAGCTTTGAAGAC 2400  
 Db 2341 AACACAGAGAGCTGGCAGAAAGAGAGAGAGATTTGAAACGTGCTGAGCTTTGAAGAC 2400  
 Qy 2401 CTTCTTGTCTTTCGCTACCAAGTGGCCAAAGGCATGGAATTCCTGAGTTCAGAGTCTGT 2460  
 Db 2401 CTTCTTGTCTTTCGCTACCAAGTGGCCAAAGGCATGGAATTCCTGAGTTCAGAGTCTGT 2460  
 Qy 2461 GTCCACAGAGACCTGGCAGCCAGGAAATGTGTGGTCAACCCACGGGAAGGTGGTGAAGATC 2520  
 Db 2461 GTCCACAGAGACCTGGCAGCCAGGAAATGTGTGGTCAACCCACGGGAAGGTGGTGAAGATC 2520  
 Qy 2521 TGTGACTTTTGGACTGGCCCGAGACATCTCTGAGGACTCCAGCTACGCTGTCAGGGGCAAC 2580  
 Db 2521 TGTGACTTTTGGACTGGCCCGAGACATCTCTGAGGACTCCAGCTACGCTGTCAGGGGCAAC 2580  
 Qy 2581 GCACGCTGCCGCTGAAGTGGATGGCACCCGAGACTTATTTGAAGGGATCTACACAATC 2640  
 Db 2581 GCACGCTGCCGCTGAAGTGGATGGCACCCGAGACTTATTTGAAGGGATCTACACAATC 2640  
 Qy 2641 AAGAGTGACGTCTGCTTACCGCATCTCTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700  
 Db 2641 AAGAGTGACGTCTGCTTACCGCATCTCTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700  
 Qy 2701 CCTTACCTCTGGCATCTCTGTCGACGCTTAACTTCTATAAATCTGATTCAGAGTGGATTTAA 2760  
 Db 2701 CCTTACCTCTGGCATCTCTGTCGACGCTTAACTTCTATAAATCTGATTCAGAGTGGATTTAA 2760  
 Qy 2761 ATGAGCAGCCATCTATGTCACAGAGGATATCTTGTGATGCAATCTCTGCTGGCT 2820  
 Db 2761 ATGAGCAGCCATCTATGTCACAGAGGATATCTTGTGATGCAATCTCTGCTGGCT 2820  
 Qy 2821 TTTGACTCAAGGAGCGGCTTCCCTCCCAACCTGACTTCTATTTTAGGATGTGAGCTG 2880  
 Db 2821 TTTGACTCAAGGAGCGGCTTCCCTCCCAACCTGACTTCTATTTTAGGATGTGAGCTG 2880  
 Qy 2881 GCAGAGCAGAAAGCATGTATCAGAAACATCTCATCTTACCAAAACAGCGGCCCT 2940  
 Db 2881 GCAGAGCAGAAAGCATGTATCAGAAACATCTCATCTTACCAAAACAGCGGCCCT 2940  
 Qy 2941 CAGCAGAGAGCGGCTCAGAGCCAGTCCGACAGAGCCAGGCTGAAATTCACAGAGAA 3000  
 Db 2941 CAGCAGAGAGCGGCTCAGAGCCAGTCCGACAGAGCCAGGCTGAAATTCACAGAGAA 3000  
 Qy 3001 AGAAGTTAGCAGAGCGGCTTGGACCCGCGCTTGGACCCGCGCTTGGACCCGCGCTTGGAC 3060  
 Db 3001 AGAAGTTAGCAGAGCGGCTTGGACCCGCGCTTGGACCCGCGCTTGGACCCGCGCTTGGAC 3060  
 Qy 3061 AGATTAGCTCTCTGAGGAGCGGCTTACAGCGCTTGTCTGCTGAGACTTTTCTCT 3120  
 Db 3061 AGATTAGCTCTCTGAGGAGCGGCTTACAGCGCTTGTCTGCTGAGACTTTTCTCTCT 3120  
 Qy 3121 AGATGCTGTCTGCCATTAATCTCAAGTGAATTTCTATAAAATCAAACTCTCTCGCAGAG 3180  
 Db 3121 AGATGCTGTCTGCCATTAATCTCAAGTGAATTTCTATAAAATCAAACTCTCTCGCAGAG 3180  
 Qy 3181 GCGGAGAGCCAAATAAGAGCTTGTGGTGGCCGCTTACCTGAGCGGCTTCCCTGGGCGCTTCCACG 3240  
 Db 3181 GCGGAGAGCCAAATAAGAGCTTGTGGTGGCCGCTTACCTGAGCGGCTTCCCTGGGCGCTTCCACG 3240  
 Qy 3241 AGCTTGGGAGGAGCCATGATCTGAAATATAGTATATTTTGTAAATPACGTGAACAA 3300  
 Db 3241 AGCTTGGGAGGAGCCATGATCTGAAATATAGTATATTTTGTAAATPACGTGAACAA 3300  
 Qy 3301 ACCAAACCGTTTTTGTCTAGGAGGAAAGCTAAATATGATTTTTTAAATAATCTATGTTTTAA 3360  
 Db 3301 ACCAAACCGTTTTTGTCTAGGAGGAAAGCTAAATATGATTTTTTAAATAATCTATGTTTTAA 3360



721 AGATGCTGTGCTAGAAATGCACTGGGCGCGAATGACCAAGCTGTTACCATAGATCTA 780  
721 AGATGCTGTGCTAGAAATGCACTGGGCGCGAATGACCAAGCTGTTACCATAGATCTA 780  
781 AACCAGGCTCTCAGAGCACACTGCCCGCTTATCTGAAAGTGGGGGAACCCCTGTGG 840  
781 AACCAGGCTCTCAGAGCACACTGCCCGCTTATCTGAAAGTGGGGGAACCCCTGTGG 840  
841 ATCAGGTGTAAGGCCATCCAATGTGAACCATGGAATTCGGGCTCACTGGGAGCTGGAAGAC 900  
841 ATCAGGTGTAAGGCCATCCAATGTGAACCATGGAATTCGGGCTCACTGGGAGCTGGAAGAC 900  
901 AAAGCCCTGAGAGGCGAGCTACTTTGAGATGAGTACTACTCCACAAACAGAGACCATG 960  
901 AAAGCCCTGAGAGGCGAGCTACTTTGAGATGAGTACTACTCCACAAACAGAGACCATG 960  
961 ATTTCGGATTCCTTGGCCTTTGTGCTTCCTGGGGAAGGACGACACCGGATATTACACC 1020  
961 ATTTCGGATTCCTTGGCCTTTGTGCTTCCTGGGGAAGGACGACACCGGATATTACACC 1020  
1021 TGCTCTTCCTCAAGACACCCAGCCAGTCAGCGTTGGTGACCATCTTAGAAAAAGGTTT 1080  
1021 TGCTCTTCCTCAAGACACCCAGCCAGTCAGCGTTGGTGACCATCTTAGAAAAAGGTTT 1080  
1081 ATAAACGCTACCACTCGCAAGAGAGATGAATAATTGACCCGTAGCAAAAGTTCTGCTTC 1140  
1081 ATAAACGCTACCACTCGCAAGAGAGATGAATAATTGACCCGTAGCAAAAGTTCTGCTTC 1140  
1141 TCAGTCAGGTTTAAAGCGTACCCAGCAATCCGATGCGACGTGATCTCTCAAGCCCTCA 1200  
1141 TCAGTCAGGTTTAAAGCGTACCCAGCAATCCGATGCGACGTGATCTCTCAAGCCCTCA 1200  
1201 TTTCTTGTGAACAGAGAGGCTGGAGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260  
1201 TTTCTTGTGAACAGAGAGGCTGGAGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260  
1261 AAGAACAGCCAGGAGAGTACATATTTCTATGACAAAATGATGAGCCCGAGTTTCAACAAA 1320  
1261 AAGAACAGCCAGGAGAGTACATATTTCTATGACAAAATGATGAGCCCGAGTTTCAACAAA 1320  
1321 ATGTTACGCTGAATATGAAGAAAGAACTCAAGTGTAGCAAAATGCTCAGCCAGCCAG 1380  
1321 ATGTTACGCTGAATATGAAGAAAGAACTCAAGTGTAGCAAAATGCTCAGCCAGCCAG 1380  
1381 CGCTCTGTTCTCTGATGGCTACCCGCTACCTCTTGACCTGGAAGAGTGTTCGGAC 1440  
1381 CGCTCTGTTCTCTGATGGCTACCCGCTACCTCTTGACCTGGAAGAGTGTTCGGAC 1440  
1441 AAATCTCCAAATTGACGAGGAAATCCCAAGAGGATTTGGAATAAAGGCTAACAGA 1500  
1441 AAATCTCCAAATTGACGAGGAAATCCCAAGAGGATTTGGAATAAAGGCTAACAGA 1500  
1501 AAAGTGTGTCAGTGGGTGTCGAGAGTACTTAAATATGAGTGAGCCCGGGAAGGG 1560  
1501 AAAGTGTGTCAGTGGGTGTCGAGAGTACTTAAATATGAGTGAGCCCGGGAAGGG 1560  
1561 CTTCTGGTCAAAATGCTGTCGTAACAAATCTATGGGCAAGCTTTGGAAACCATCTTTTAA 1620  
1561 CTTCTGGTCAAAATGCTGTCGTAACAAATCTATGGGCAAGCTTTGGAAACCATCTTTTAA 1620  
1621 AACTCAGAGGCCCCCTTCCCTTTTCAATCCAAAGACCAATCTCTTCTATGCGACCATTTGG 1680  
1621 AACTCAGAGGCCCCCTTCCCTTTTCAATCCAAAGACCAATCTCTTCTATGCGACCATTTGG 1680  
1681 CTTCTGCTCCCCCTTCAATGTTGTTCTCATATGTTGTTGATCTGCGCAAAATACAAAAGCAA 1740  
1681 CTTCTGCTCCCCCTTCAATGTTGTTCTCATATGTTGTTGATCTGCGCAAAATACAAAAGCAA 1740  
1741 TTTAGGTACAGAGTCAGTCGAGATGATCCAGGTGACTGGCCCCCTGGATACAGAGTAC 1800  
1741 TTTAGGTACAGAGTCAGTCGAGATGATCCAGGTGACTGGCCCCCTGGATACAGAGTAC 1800  
1801 TTTCTAGTTGACTTTCAGGAGCTATGAATATGACCTTAAAGTGGGAGTTCCCGAGAGAGAAC 1860

1801 TTCTAGTTGACTTTCAGGAGCTATGAATATGACCTTAAAGTGGGAGTTCCCGAGAGAGAC 1860  
1861 TTAGAGTTTGGGAAGCTCTCGGCTCTGGGCTTTCGGGAGGCTGATGAACCGCACCGGCC 1920  
1861 TTAGAGTTTGGGAAGCTCTCGGCTCTGGGCTTTCGGGAGGCTGATGAACCGCACCGGCC 1920  
1921 TATGGCATTAGTAAAAACGGGAGTCTCAATTCAGGTGCGCTGGAAGATGCTAAAAAGAGAA 1980  
1921 TATGGCATTAGTAAAAACGGGAGTCTCAATTCAGGTGCGCTGGAAGATGCTAAAAAGAGAA 1980  
1981 GCTGACAGCTGTGAAAAAGAAAGCTCTCATGTGCGAGCTCAAAAATGATGACCCACTGGGA 2040  
1981 GCTGACAGCTGTGAAAAAGAAAGCTCTCATGTGCGAGCTCAAAAATGATGACCCACTGGGA 2040  
2041 CACCATGACAAACATCGTGAATCTGCTGGGGCATGACACATGTCAGGGCCAGTTGACTTG 2100  
2041 CACCATGACAAACATCGTGAATCTGCTGGGGCATGACACATGTCAGGGCCAGTTGACTTG 2100  
2101 ATTTTGAATATTTGCTGCTATGCTGACCTCTCTCAACTACCTTAAGAAGTAAAGAGAGAG 2160  
2101 ATTTTGAATATTTGCTGCTATGCTGACCTCTCTCAACTACCTTAAGAAGTAAAGAGAGAG 2160  
2161 TTTTCAAGGACATGGAACAGAGATTTTAAAGAAATATTTCAAGTTCTTACCTACTTTC 2220  
2161 TTTTCAAGGACATGGAACAGAGATTTTAAAGAAATATTTCAAGTTCTTACCTACTTTC 2220  
2221 CAGGCAATTTCAAAATTCAGCATGCTGTTTCAAGAGAGTTCAGTTACACCCGCCCTTG 2280  
2221 CAGGCAATTTCAAAATTCAGCATGCTGTTTCAAGAGAGTTCAGTTACACCCGCCCTTG 2280  
2281 GATCAGCTCTCAGGCTTCAATGSGAAATTCATTTCAATCTGAAGATGAGATTGAAATGAA 2340  
2281 GATCAGCTCTCAGGCTTCAATGSGAAATTCATTTCAATCTGAAGATGAGATTGAAATGAA 2340  
2341 AACAGAGAGGCTGGCAGAGAGAGAGAGAGATTTGAAACGCTGCTGACGTTTGAAGAC 2400  
2341 AACAGAGAGGCTGGCAGAGAGAGAGAGAGAGATTTGAAACGCTGCTGACGTTTGAAGAC 2400  
2401 CTCCTTTGCTTGGCTACCAAGTGCGCAAGGATGGAATTCCTGAGTTTCAAGTCTGAT 2460  
2401 CTCCTTTGCTTGGCTACCAAGTGCGCAAGGATGGAATTCCTGAGTTTCAAGTCTGAT 2460  
2461 GTCCACAGAGAGCTGGCAGCAGGAAATCTGTTGTCACCCAGGAGGCTGAGAGATC 2520  
2461 GTCCACAGAGAGCTGGCAGCAGGAAATCTGTTGTCACCCAGGAGGCTGAGAGATC 2520  
2521 TGTGACTTTGGACTGGCCCGAGACATCTGAGCGACTCCAGCTACGTTCCAGGGGCAAC 2580  
2521 TGTGACTTTGGACTGGCCCGAGACATCTGAGCGACTCCAGCTACGTTCCAGGGGCAAC 2580  
2581 GCACGCTGCTGGCTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGAGTCTACACAATC 2640  
2581 GCACGCTGCTGGCTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGAGTCTACACAATC 2640  
2641 AAGAGTGAGCTGCTGCTTCAAGGATTCCTTCTCTGAGAGATTTTCACTGGGCTGGAAC 2700  
2641 AAGAGTGAGCTGCTGCTTCAAGGATTCCTTCTCTGAGAGATTTTCACTGGGCTGGAAC 2700  
2701 CTTTACCTGGCATTTCTGTCGACGCTAACTTCTATAAACTGATTTCAGAGTGGATTAAA 2760  
2701 CTTTACCTGGCATTTCTGTCGACGCTAACTTCTATAAACTGATTTCAGAGTGGATTAAA 2760  
2761 ATGGAGAGCCATTTATGCCACAGAGGAGATATCTTTGTAATGCAATCTCTCGGCT 2820  
2761 ATGGAGAGCCATTTATGCCACAGAGGAGATATCTTTGTAATGCAATCTCTCGGCT 2820  
2821 TTTGACTCAAGAGAGCGGCAATCTTCCCAACCTGACTTCAATTTTAGGATGTCAGCTG 2880  
2821 TTTGACTCAAGAGAGCGGCAATCTTCCCAACCTGACTTCAATTTTAGGATGTCAGCTG 2880  
2881 GCAGAGGCAAGAGAGATGATCAAGACATCCATCCATCTACCAAAACAGGGGCGCCCT 2940



601	CGGGAGCCACTGTGGAGTGGGTCTCTGAGCTCCACAGGGAAAGCTGTAAAGAGAA	660	1681	CTCTGTCCTCCCTTCATTGTTGTTCTCAITGTTGTTGATCTGCCACAATAACAAAAGCAA	1740
661	GGCCCTGCTGTTGTAGAAAGAGGAAAGAGTACTTCAAGAGTGTTCGGAAACAGACATC	720	1741	TTTAGTACAGAGTACAGCTGCGAGATGATCCAGTGACTGCGCCCTCGGATACACGAGTAC	1800
661	GGCCCTGCTGTTGTAGAAAGAGGAAAGAGTACTTCAAGAGTGTTCGGAAACAGACATC	720	1741	TTTAGTACAGAGTACAGCTGCGAGATGATCCAGTGACTGCGCCCTCGGATACACGAGTAC	1800
721	AGATGCTGTGTAGAAATGCACTGGGCGCGAAATGCACCAAGCTGTTCACATAGATCTTA	780	1801	TTCTACGTTGACTTTCAGGGACTATGAATATGACCTTAAAGTGGGAGTTCCCGAGAGAGAA	1860
721	AGATGCTGTGTAGAAATGCACTGGGCGCGAAATGCACCAAGCTGTTCACATAGATCTTA	780	1801	TTCTACGTTGACTTTCAGGGACTATGAATATGACCTTAAAGTGGGAGTTCCCGAGAGAGAA	1860
781	AAACAGGCTCTCAGAGCACTGCCCCAGTTATTCCTGAAAGTGGGGAAACCTTGTGG	840	1861	TTAGAGTTTGGGAAGTCTCTGGGGTCTGGCGCTTTCGGGAGGGTGAAGACCCACGGCC	1920
781	AAACAGGCTCTCAGAGCACTGCCCCAGTTATTCCTGAAAGTGGGGAAACCTTGTGG	840	1861	TTAGAGTTTGGGAAGTCTCTGGGGTCTGGCGCTTTCGGGAGGGTGAAGACCCACGGCC	1920
841	ATCAGGTGAAGCCATCCATGTGAACCATGGATTGGGGTCACTCGGGAGCTGGAGAGC	900	1921	TATGGCATTTAGTAAACCGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTTAAAGAGAGAA	1980
841	ATCAGGTGAAGCCATCCATGTGAACCATGGATTGGGGTCACTCGGGAGCTGGAGAGC	900	1921	TATGGCATTTAGTAAACCGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTTAAAGAGAGAA	1980
901	AAAGCCCTGGAGGAGGAGCTACTTTGAGATGAGTACTTCTCCACAAACAGGACCATG	960	1981	GCTGACAGCTGTGAAAAGAAAGCTCTCATGTGGAGCTCAAAAATGATGACCCACCTGGGA	2040
901	AAAGCCCTGGAGGAGGAGCTACTTTGAGATGAGTACTTCTCCACAAACAGGACCATG	960	1981	GCTGACAGCTGTGAAAAGAAAGCTCTCATGTGGAGCTCAAAAATGATGACCCACCTGGGA	2040
961	ATTTCGGATCTCTTGGCCTTGTGTCTTCCTGGGAAAGGAAAGACACCGGATATTACAC	1020	2041	CACCATGACAAACATCGTGAATCTGCTGGGGCATGCACTGTGAGGGCCAGTGACTTGTG	2100
961	ATTTCGGATCTCTTGGCCTTGTGTCTTCCTGGGAAAGGAAAGACACCGGATATTACAC	1020	2041	CACCATGACAAACATCGTGAATCTGCTGGGGCATGCACTGTGAGGGCCAGTGACTTGTG	2100
1021	TGCTCTTCCTCAAGACACCCAGCAGTACCGTGTGGTACCATCTCTAGAAAAGGGTTT	1080	2101	ATTTTGAATATTTGCTATGTTGCTATGTTGCTCTCTCAACTACCTAAGAAAGTAAAGAGAG	2160
1021	TGCTCTTCCTCAAGACACCCAGCAGTACCGTGTGGTACCATCTCTAGAAAAGGGTTT	1080	2101	ATTTTGAATATTTGCTATGTTGCTCTCTCTCAACTACCTAAGAAAGTAAAGAGAG	2160
1081	ATAAACGCTACAGCTCGAAGAGAGTATGAAATGACCCGTAGCAAAAGTCTCTGCTTC	1140	2161	TTTTCACAGGACATGACAGAGATTTTAAAGAAACATAATTTTACCTTACTTACTTCTTC	2220
1081	ATAAACGCTACAGCTCGAAGAGAGTATGAAATGACCCGTAGCAAAAGTCTCTGCTTC	1140	2161	TTTTCACAGGACATGACAGAGATTTTAAAGAAACATAATTTTACCTTACTTACTTCTTC	2220
1141	TCAGTCAGGTTTAAAGCGTACCCAGAAATCCGATGCGACGTGGAATCTCTCAAGCCTCA	1200	2221	CAGGCACATTTCAAATTCAGCATGCTGTTTCAAGAAAGTTCAGTTACACCCGCCCTTG	2280
1141	TCAGTCAGGTTTAAAGCGTACCCAGAAATCCGATGCGACGTGGAATCTCTCAAGCCTCA	1200	2221	CAGGCACATTTCAAATTCAGCATGCTGTTTCAAGAAAGTTCAGTTACACCCGCCCTTG	2280
1201	TTTCTTTGTAACAGAGAGGCTGGAGATGGGTACAGCATCTAAATTTTGGGATCAT	1260	2281	GATCAGCTCTCAGGGTTCATATGGGAAATTCATTTCTTGAAGATGAGATTGAAATAGAA	2340
1201	TTTCTTTGTAACAGAGAGGCTGGAGATGGGTACAGCATCTAAATTTTGGGATCAT	1260	2281	GATCAGCTCTCAGGGTTCATATGGGAAATTCATTTCTTGAAGATGAGATTGAAATAGAA	2340
1261	AAGAACAGCCAGAGAGTACATATTTCTATGCAAAATGATGACCCAGTTCACCAAA	1320	2341	AACAGAAAGAGCTGGCAG	2400
1261	AAGAACAGCCAGAGAGTACATATTTCTATGCAAAATGATGACCCAGTTCACCAAA	1320	2341	AACAGAAAGAGCTGGCAG	2400
1321	ATGTTACGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCTCAGCCAGCCAG	1380	2401	CTCCTTTCTTTGCTTACCAAGTGGCCAAAGGATGGAATTTGAACTGCTGAGCTTTGAGAC	2460
1321	ATGTTACGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCTCAGCCAGCCAG	1380	2401	CTCCTTTCTTTGCTTACCAAGTGGCCAAAGGATGGAATTTGAACTGCTGAGCTTTGAGAC	2460
1381	CGCTCTGTTCTCTGATGGCTACCCGTACCTCTTGGACCTGGAAGAGTGTTCGAC	1440	2461	GTCCACAGAGACCTGGCAGCCAGGAATGTGTTGGTCAACCAAGGAGTGTGTAAGATC	2520
1381	CGCTCTGTTCTCTGATGGCTACCCGTACCTCTTGGACCTGGAAGAGTGTTCGAC	1440	2461	GTCCACAGAGACCTGGCAGCCAGGAATGTGTTGGTCAACCAAGGAGTGTGTAAGATC	2520
1441	AAATCTCCAAATGCAACGAGAAATCCAGAGAGGTTTGAATAAAGGCTAACAGA	1500	2521	TGTGACTTTGAGCTGGCCGAGACATCTCTGAGGACTCCAGCTACCTGCTGAGGAGCAAC	2580
1441	AAATCTCCAAATGCAACGAGAAATCCAGAGAGGTTTGAATAAAGGCTAACAGA	1500	2521	TGTGACTTTGAGCTGGCCGAGACATCTCTGAGGACTCCAGCTACCTGCTGAGGAGCAAC	2580
1501	AAAGTGTGTCAGTGGTGTGAGAGAGTCTCTAAATATGAGTGGAGGAGGAGGAGG	1560	2581	GCACGGCTGCGGCTCAAGTGGATGCGACCCGAGAGCTTATTTGAAGGAGTCTACACAATC	2640
1501	AAAGTGTGTCAGTGGTGTGAGAGAGTCTCTAAATATGAGTGGAGGAGGAGGAGG	1560	2581	GCACGGCTGCGGCTCAAGTGGATGCGACCCGAGAGCTTATTTGAAGGAGTCTACACAATC	2640
1561	CTTCTGGTCAAAATGCTGCGTACAAATCTATGGGCACTGCTTGGGAAACCATCTTTT	1620	2641	AAGAGTGAAGCTGCTGGTCTTACCGGATCTCTTCTCTGGAGATATTTTCACTGGGTGAAC	2700
1561	CTTCTGGTCAAAATGCTGCGTACAAATCTATGGGCACTGCTTGGGAAACCATCTTTT	1620	2641	AAGAGTGAAGCTGCTGGTCTTACCGGATCTCTTCTCTGGAGATATTTTCACTGGGTGAAC	2700
1621	AACCTCAGGAGGCTTCTCTTCTATCCAGACACATCTCTTCTATGCGACCATTTGG	1680	2701	CCTTACCTTGGCATCTCTGCGACGCTTCTTCTATAAACTGATTCAGAGTGGATTAAA	2760
1621	AACCTCAGGAGGCTTCTCTTCTATCCAGACACATCTCTTCTATGCGACCATTTGG	1680	2701	CCTTACCTTGGCATCTCTGCGACGCTTCTTCTATAAACTGATTCAGAGTGGATTAAA	2760
1681	CTCTGCTCTCCCTTCATTGTTGTTCTCAITGTTGTTGATCTGCCACAATAACAAAAGCAA	1740	2761	ATGGAGCAGCCATTCTATGCCACAGAGGATATCTTTGTAATGCAATCTCTGCTGGGCT	2820
1681	CTCTGCTCTCCCTTCATTGTTGTTCTCAITGTTGTTGATCTGCCACAATAACAAAAGCAA	1740	2761	ATGGAGCAGCCATTCTATGCCACAGAGGATATCTTTGTAATGCAATCTCTGCTGGGCT	2820

QY 2821 TTTGACTCAAGGAAGCGGCCATCTCTCCCAACCTGACTTCATTTTGGATGTCAGCTG 2880  
 Db |||||||  
 QY 2821 TTTGACTCAAGGAAGCGGCCATCTCTCCCAACCTGACTTCATTTTGGATGTCAGCTG 2880  
 Db |||||||  
 QY 2881 GCAGAGGCAGGAAGCATGTATCAGAACATCCATCCATCTACCAAAACAGGCGGCCCT 2940  
 Db |||||||  
 QY 2881 GCAGAGGCAGGAAGCATGTATCAGAACATCCATCCATCTACCAAAACAGGCGGCCCT 2940  
 Db |||||||  
 QY 2941 CAGCAGAGGCGGGCTCAGAGCCAGTCGCGCCAGCGCCAGGTGAAGATTCAGAGAA 3000  
 Db |||||||  
 QY 2941 CAGCAGAGGCGGGCTCAGAGCCAGTCGCGCCAGCGCCAGGTGAAGATTCAGAGAA 3000  
 Db |||||||  
 QY 3001 AGAAGTTAGCAGAGGCGCTTGAGCCCGCCAGCCAGCGCTGTAGACCGCAGAGCCA 3060  
 Db |||||||  
 QY 3001 AGAAGTTAGCAGAGGCGCTTGAGCCCGCCAGCCAGCGCTGTAGACCGCAGAGCCA 3060  
 Db |||||||  
 QY 3061 AGATTAGCCTCGCTCTGAGGAAGCGCCCTACAGCGGCTTCTGCTGACTTTCTCT 3120  
 Db |||||||  
 QY 3061 AGATTAGCCTCGCTCTGAGGAAGCGCCCTACAGCGGCTTCTGCTGACTTTCTCT 3120  
 Db |||||||  
 QY 3121 AGATGCTGCTGCTTACTTCCAAAGTGACTTCTATAAAATCAAACTCTCTCGCACAG 3180  
 Db |||||||  
 QY 3121 AGATGCTGCTGCTTACTTCCAAAGTGACTTCTATAAAATCAAACTCTCTCGCACAG 3180  
 Db |||||||  
 QY 3181 GCGGAGAGCCAAATATGAGACTTGTGGTGAGCCCGCCCTACCTCGGGGCTTTCCAG 3240  
 Db |||||||  
 QY 3181 GCGGAGAGCCAAATATGAGACTTGTGGTGAGCCCGCCCTACCTCGGGGCTTTCCAG 3240  
 Db |||||||  
 QY 3241 ACCTTGAGGGGAAAGCCATGTATCTGAATATAGTATATCTTGTAAATACGTGAACAA 3300  
 Db |||||||  
 QY 3241 ACCTTGAGGGGAAAGCCATGTATCTGAATATAGTATATCTTGTAAATACGTGAACAA 3300  
 Db |||||||  
 QY 3301 ACCAAACCGCTTTTGTCTAAGGAAGCTAAATATGATTTTAAATCTATGTTTAA 3360  
 Db |||||||  
 QY 3301 ACCAAACCGCTTTTGTCTAAGGAAGCTAAATATGATTTTAAATCTATGTTTAA 3360  
 Db |||||||  
 QY 3361 AATACTATGTAATTTTCTCATCTATTAGTGATATATTTTATGGATGGAATAACTTTC 3420  
 Db |||||||  
 QY 3361 AATACTATGTAATTTTCTCATCTATTAGTGATATATTTTATGGATGGAATAACTTTC 3420  
 Db |||||||  
 QY 3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3453  
 Db |||||||  
 QY 3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3453  
 Db |||||||

## RESULT 7

ID AAQ35249 standard; cDNA; 3453 BP.  
 XX

AC AAQ35249;

XX

XX

DT 25-MAR-2003 (updated)

XX 25-JUN-1993 (first entry)

XX Murine flk-2 coding sequence.

KW Murine; receptor; protein; tyrosine kinase; pTK; primitive; mammalian;  
 KW hematopoietic cell; pHC; mature; mHC; fetal; liver kinase 2; flk-2;  
 KW liver; spleen; thymus; adult; brain; marrow; thymocyte; subset;  
 KW multipotential; T-lymphoid; lineage; ss.  
 XX

OS Mus musculus.

XX

XX Key Location/Qualifiers

FT CDS 31..3009

FT /\*tag= a

FT misc\_RNA 31..111

FT /\*tag= b

XX /note= "Hydrophobic leader"

XX WO9300349-A1.

XX

PD 07-JAN-1993.

XX

PF 26-JUN-1992; 92WO-US05401.

XX

PR 28-JUN-1991; 91US-0728913.

PR 15-NOV-1991; 91US-0793065.

PR 24-DEC-1991; 91US-0813593.

PR 02-APR-1992; 92WO-US02750.

XX (UYPR-) UNIV PRINCETON.

PA

XX Lemischka IR;

XX

XX WPI; 1993-036323/04.

DR P-PSDB; AAR31375.

XX

PT Nucleic acid encoding receptor protein tyrosine kinase - allows

PT development of ligands to stimulate proliferation and/or

PT differentiation of mammalian haematopoietic stem cells

XX Claim 5; Fig 1a; 78pp; English.

PS

XX This sequence encodes a murine receptor protein tyrosine kinase which

CC belongs to a new functional class of protein tyrosine kinases (pTKs).

CC pTKs in this class are expressed in primitive mammalian hematopoietic

CC (pHC) cells but not in mature hematopoietic cells (mHC). The protein

CC encoded by this sequence is an example of a receptor pTK and is called

CC fetal liver kinase 2 (flk-2). flk-2 is expressed in fetal liver,

CC spleen and thymus, and adult brain and marrow. Expression of flk-2

CC mRNA occurs in the most primitive thymocyte subset, which is believed

CC to be uncommitted. Therefore, thymocytes expressing flk-2 may be

CC multipotential. flk-2 is the first receptor tyrosine kinase known to

CC be expressed in the T-lymphoid lineage.

CC (Updated on 25-MAR-2003 to correct PN field.)

XX

SQ Sequence 3453 BP; 947 A; 821 C; 876 G; 809 T; 0 other;

Query Match 100.0%; Score 3451.4; DB 14; Length 3453;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3452; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCGGCTGCTACCGCGCTCCGGAGGCCATCGGGCGTTGGCGAGCGGACCGG 60

Db |||||||

QY 1 GCGGCTGCTACCGCGCTCCGGAGGCCATCGGGCGTTGGCGAGCGGACCGG 60

Db |||||||

QY 61 CGGCTGCTGCTGCTGTTGTTTGTGCTAGTAATGTTTGTGAGACCGTTACAAACCAAGAC 120

Db |||||||

QY 61 CGGCTGCTGCTGCTGTTGTTTGTGCTAGTAATGTTTGTGAGACCGTTACAAACCAAGAC 120

Db |||||||

QY 121 CTGCTGTGATCAAGTGTTGTTTAAATCAGTCATGAGAACAAATGGCTCATCGCGGGAAG 180

Db |||||||

QY 121 CTGCTGTGATCAAGTGTTGTTTAAATCAGTCATGAGAACAAATGGCTCATCGCGGGAAG 180

Db |||||||

QY 181 CCATCATCTGACCAATGGTGGAGGATCCCAAGACCTCCAGTGTACCCGAGGCGC 240

Db |||||||

QY 181 CCATCATCTGACCAATGGTGGAGGATCCCAAGACCTCCAGTGTACCCGAGGCGC 240

Db |||||||

QY 241 CAGAGTGAAGGCGGATATGAGCGGCCACCGTGGAGGTGGCGGAGTCTGGTCCATC 300

Db |||||||

QY 241 CAGAGTGAAGGCGGATATGAGCGGCCACCGTGGAGGTGGCGGAGTCTGGTCCATC 300

Db |||||||

QY 301 ACCCTGCAAGTGACGCTCGCCACCCAGGGGACCTTTCTCCCTCTGGTCTTTAAGCAC 360

Db |||||||

QY 301 ACCCTGCAAGTGACGCTCGCCACCCAGGGGACCTTTCTCCCTCTGGTCTTTAAGCAC 360

Db |||||||

QY 361 AGTCCCTGGGTGCGCAGCGCACTTTGATTTTACAAACAGAGGAATCGTTTCCATGCC 420

Db |||||||

QY 361 AGTCCCTGGGTGCGCAGCGCACTTTGATTTTACAAACAGAGGAATCGTTTCCATGCC 420

Db |||||||

QY 421 ATCTTGAACTGACAGAGACCCAGGAGGAGATACCTACTCCATATTCAGAGGCAAGC 480

Db |||||||

QY 421 ATCTTGAACTGACAGAGACCCAGGAGGAGATACCTACTCCATATTCAGAGGCAAGC 480

Db |||||||



481 GCCAATACACAGTACTGTTACAGTGAATGTAAGAGATACACAGCTGTATGTCTAAGG 540  
 481 GCCAATACACAGTACTGTTACAGTGAATGTAAGAGATACACAGCTGTATGTCTAAGG 540  
 541 AGACCTTACTTTAGGAAGATGGAACACAGGATGCACTGTCTGTCTCTCCAGGGTGT 600  
 541 AGACCTTACTTTAGGAAGATGGAACACAGGATGCACTGTCTGTCTCTCCAGGGTGT 600  
 601 CCGAGGCCATCTGTGGAGTGGTCTCTGACGTCCTCCACAGGAAAGCTGTAAAGAGAA 660  
 601 CCGAGGCCATCTGTGGAGTGGTCTCTGACGTCCTCCACAGGAAAGCTGTAAAGAGAA 660  
 661 GGCCCTCTGTTTTCAGAAAGGAGAAAGTACTTCAATGATGTTGTTCCGGAACAGACATC 720  
 661 GGCCCTCTGTTTTCAGAAAGGAGAAAGTACTTCAATGATGTTGTTCCGGAACAGACATC 720  
 721 AGATGCTGTGTAGAAATGCACTGGGCGCGAATGCAACCAAGCTGTTCCACATAGATCTA 780  
 721 AGATGCTGTGTAGAAATGCACTGGGCGCGAATGCAACCAAGCTGTTCCACATAGATCTA 780  
 781 AACCAAGCTCTCTAGAGACACTGCCCCAGTTATTCCTGAAAGTGGGGGAAACCCCTGTGG 840  
 781 AACCAAGCTCTCTAGAGACACTGCCCCAGTTATTCCTGAAAGTGGGGGAAACCCCTGTGG 840  
 841 ATCAGGTGTAAAGCCATCCATGTGAACCATGGAATTCGGGCTCACCTGGAGCTGGAAGAC 900  
 841 ATCAGGTGTAAAGCCATCCATGTGAACCATGGAATTCGGGCTCACCTGGAGCTGGAAGAC 900  
 901 AAAGCCCTGGAGGAGGAGTACTTTCAGATGAGTACTTCCACAAACAGGACCATG 960  
 901 AAAGCCCTGGAGGAGGAGTACTTTCAGATGAGTACTTCCACAAACAGGACCATG 960  
 961 ATTCCGATTTCTTGGCCCTTTGTGTTCTCCGTTGGGAAGAACACACCGGATATTACACC 1020  
 961 ATTCCGATTTCTTGGCCCTTTGTGTTCTCCGTTGGGAAGAACACACCGGATATTACACC 1020  
 1021 TGCTCTTCTCAAGCAACCCAGCAGTACAGGTTGGTGACCATCTGTGAAGAAAGGGT 1080  
 1021 TGCTCTTCTCAAGCAACCCAGCAGTACAGGTTGGTGACCATCTGTGAAGAAAGGGT 1080  
 1081 ATAAACGCTACAGCTCCGACAGAGATGAAATTTGACCCCTGACGAAAGTTCTGCTTC 1140  
 1081 ATAAACGCTACAGCTCCGACAGAGATGAAATTTGACCCCTGACGAAAGTTCTGCTTC 1140  
 1141 TCAGTCAGGTTTAAAGCGTACCCAGATCCGATGCACTGGATCTTCTCTCAAGCCTCA 1200  
 1141 TCAGTCAGGTTTAAAGCGTACCCAGATCCGATGCACTGGATCTTCTCTCAAGCCTCA 1200  
 1201 TTTCTCTGTGAACAGAGAGGCTGGAGGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260  
 1201 TTTCTCTGTGAACAGAGAGGCTGGAGGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260  
 1261 AAGAACAGCAGGAGATGATATTTCTATGAGAAATGATGACCCAGTTCCACAA 1320  
 1261 AAGAACAGCAGGAGATGATATTTCTATGAGAAATGATGACCCAGTTCCACAA 1320  
 1321 ATGTTTCAGCTCAATATAGAAAGAACCTCAAGTCTAGCAATTCGCTCAGCCAGCAG 1380  
 1321 ATGTTTCAGCTCAATATAGAAAGAACCTCAAGTCTAGCAATTCGCTCAGCCAGCAG 1380  
 1381 GGCTCTGTTCTCTGATGGCTACCCGCTACCTCTTGGACCTGGGAAGTGTTCGGAC 1440  
 1381 GGCTCTGTTCTCTGATGGCTACCCGCTACCTCTTGGACCTGGGAAGTGTTCGGAC 1440  
 1441 AAATCTCCAAATGCAAGGAGAAATCCCGAAGGAGTTTGGATATAAAGGCTAACAGA 1500  
 1441 AAATCTCCAAATGCAAGGAGAAATCCCGAAGGAGTTTGGATATAAAGGCTAACAGA 1500  
 1501 AAAGTGTGTTGGCAGTGGTGTGAGCAGTACTTAAATATGATGAGGCGGGAAGGG 1560  
 1501 AAAGTGTGTTGGCAGTGGTGTGAGCAGTACTTAAATATGATGAGGCGGGAAGGG 1560  
 1561 CTTCTGGTCAATGTCTGCTGCTACAATTTCTATGGGACGCTTTTGGCAACCATCTTTT 1620

1561 CTTCTGGTCAATGTCTGCTGCTACAATTTCTATGGGACGCTTTGCAAAACCATCTTTT 1620  
 1621 AACTCACAGGCCCCCTTCCCTTTTCATCAACAGCAACATCTCTCTTAAGCAACCATTTGG 1680  
 1621 AACTCACAGGCCCCCTTCCCTTTTCATCAACAGCAACATCTCTCTTAAGCAACCATTTGG 1680  
 1681 CTTCTGCTCCCTTTCATTTGTTTCTCATTTGTTGATCTGCCCAAAATACAAAAGCAA 1740  
 1681 CTTCTGCTCCCTTTCATTTGTTTCTCATTTGTTGATCTGCCCAAAATACAAAAGCAA 1740  
 1741 TTTAGGTACGAGATCAGTCTGAGTATGATCCAGGTGACTGGCCCCCTGGATTAACAGTAC 1800  
 1741 TTTAGGTACGAGATCAGTCTGAGTATGATCCAGGTGACTGGCCCCCTGGATTAACAGTAC 1800  
 1801 TTTACGTTGACTTTCAGGACTATGATGATGATGATGATGATGATGATGATGATGATGAT 1860  
 1801 TTTACGTTGACTTTCAGGACTATGATGATGATGATGATGATGATGATGATGATGATGAT 1860  
 1861 TTAGGATTTGGGAAGGCTCTGGGCTCTGGGCTCTGGGCTCTGGGCTCTGGGCTCTGGGCT 1920  
 1861 TTAGGATTTGGGAAGGCTCTGGGCTCTGGGCTCTGGGCTCTGGGCTCTGGGCTCTGGGCT 1920  
 1921 TATGCAATTTAGTAAACCGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAAGAGAA 1980  
 1921 TATGCAATTTAGTAAACCGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAAGAGAA 1980  
 1981 GCTGACAGCTGTGAAAAAGAGTCTCATGTCGGAGCTCAAAATGATGATGATGATGATGAT 2040  
 1981 GCTGACAGCTGTGAAAAAGAGTCTCATGTCGGAGCTCAAAATGATGATGATGATGATGAT 2040  
 2041 CACCATGACACATCGTGAATCTGTCGGGGCATGCACACTCTCAGGCGCAGTGTACTTTC 2100  
 2041 CACCATGACACATCGTGAATCTGTCGGGGCATGCACACTCTCAGGCGCAGTGTACTTTC 2100  
 2101 ATTTTGAATATTTGTTGCTATGTTGCTCAACTCTCAACTCTCAACTCTCAACTCTCAACT 2160  
 2101 ATTTTGAATATTTGTTGCTATGTTGCTCAACTCTCAACTCTCAACTCTCAACTCTCAACT 2160  
 2161 TTTTCAAGGACATGACAGAGATTTTAAAGAACATAATTTCAAGTTCTTACCCCTTTC 2220  
 2161 TTTTCAAGGACATGACAGAGATTTTAAAGAACATAATTTCAAGTTCTTACCCCTTTC 2220  
 2221 CAGGCACATTTCAATTTCCAGCATGCTGTCCTCAGGAGAGTTCAGTTTACACCCGCTTC 2280  
 2221 CAGGCACATTTCAATTTCCAGCATGCTGTCCTCAGGAGAGTTCAGTTTACACCCGCTTC 2280  
 2281 GATCAGCTCTCAGGCTTCAATGGGAATTCATTTCAATTTCAATTTCAATTTCAATTTCAAT 2340  
 2281 GATCAGCTCTCAGGCTTCAATGGGAATTCATTTCAATTTCAATTTCAATTTCAATTTCAAT 2340  
 2341 AACCAAGAGAGCTGGCAG 2400  
 2341 AACCAAGAGAGCTGGCAG 2400  
 2401 CTTCTTTGCTTTGCTGTTTCAAGTGGCCAAAGGATGCTGCTGCTGCTGCTGCTGCTGCTG 2460  
 2401 CTTCTTTGCTTTTGGTACCAAGTGGCCAAAGGATGCTGCTGCTGCTGCTGCTGCTGCTG 2460  
 2461 GTTCCACAGAGAGCTGGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520  
 2461 GTTCCACAGAGAGCTGGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520  
 2521 TGTGATTTTGGATGCTGGCCCGGAGATCTCTGAGCAGTCTCCAGCTACGTCGTGAGGG 2580  
 2521 TGTGATTTTGGATGCTGGCCCGGAGATCTCTGAGCAGTCTCCAGCTACGTCGTGAGGG 2580  
 2581 GCAAGCTGCTGGTGAAGTGGATGGCAACCGGAGAGCTTTATTTGAAGGATCTACACATC 2640  
 2581 GCAAGCTGCTGGTGAAGTGGATGGCAACCGGAGAGCTTTATTTGAAGGATCTACACATC 2640  
 2641 AAGAGTGAAGTCTGGTCTTACGCTACGCTCTCTCTGGGAGATATTTTCACTGGGTGTGAC 2700

```

Db      2641 AAGAGTACGCTCTGGTCTCTACGGCATCCTTCTCTGGAGATATTTTCACTGGGTGTGAAC 2700
Qy      CTTTACCCTGGCAATTCCTGTGCGAGCTAACTTCTATATAAAGTATTCAGAGTGGATTTAAA 2760
Db      2701 CTTTACCCTGGCAATTCCTGTGCGAGCTAACTTCTATATAAAGTATTCAGAGTGGATTTAAA 2760
Qy      2761 ATGAGAGGCAATTCATAGCCACAGAGGAGATATCTTTGTATGCAATCTCTGGGCT 2820
Db      2761 ATGAGAGGCAATTCATAGCCACAGAGGAGATATCTTTGTATGCAATCTCTGGGCT 2820
Qy      2821 TTTGACTCAAGGAAGCGGCATCTTCCCAACCTGACTCATTTTATGAGATGCACTG 2880
Db      2821 TTTGACTCAAGGAAGCGGCATCTTCCCAACCTGACTCATTTTATGAGATGCACTG 2880
Qy      2881 GCAGAGCGAGAAAGAGCATGTATCAGAAATCATCATCATCTACCAAAACAGGGCCCT 2940
Db      2881 GCAGAGCGAGAAAGAGCATGTATCAGAAATCATCATCATCTACCAAAACAGGGCCCT 2940
Qy      2941 CAGCAGAGGCGGGCTCAGAGCCAGTGCACAGCGCCAGTGCAGGAGGATTCACAGAGAA 3000
Db      2941 CAGCAGAGGCGGGCTCAGAGCCAGTGCACAGCGCCAGTGCAGGAGGATTCACAGAGAA 3000
Qy      3001 AGAAGTTAGCGAGGAGGCTTGCACCGCCAGTGCAGGCTGTAGACCGCAGAGCCA 3060
Db      3001 AGAAGTTAGCGAGGAGGCTTGCACCGCCAGTGCAGGCTGTAGACCGCAGAGCCA 3060
Qy      3061 AGATTAGCTCGCTCTGAGGAAGCGCCCTCAGAGCGGCTTCTCTGAGACTTTCTCT 3120
Db      3061 AGATTAGCTCGCTCTGAGGAAGCGCCCTCAGAGCGGCTTCTCTGAGACTTTCTCT 3120
Qy      3121 AGATCGTGTGCCATTTACTCCAAAGTGAATCTTATATAAATCAAACTCTCTCGCACAG 3180
Db      3121 AGATCGTGTGCCATTTACTCCAAAGTGAATCTTATATAAATCAAACTCTCTCGCACAG 3180
Qy      3181 GCGGAGAGCCCAATATAGACTTGTGTGAGCGCCCTACCTCGGGGCGCTTCCAGC 3240
Db      3181 GCGGAGAGCCCAATATAGACTTGTGTGAGCGCCCTACCTCGGGGCGCTTCCAGC 3240
Qy      3241 AGCTTGAGGGAAGCCATGTATCTGAAATAGTATATTTCTGTAATACGTGAACAA 3300
Db      3241 AGCTTGAGGGAAGCCATGTATCTGAAATAGTATATTTCTGTAATACGTGAACAA 3300
Qy      3301 ACCAAACCCGTTTTTGTGTAAGGGAAGCTAAATAGTATTTTAAATCTATGTTTAA 3360
Db      3301 ACCAAACCCGTTTTTGTGTAAGGGAAGCTAAATAGTATTTTAAATCTATGTTTAA 3360
Qy      3361 AATAGTATGAACTTTTCTATCTATTTAGTGATATATTTTATGATGGAATAAACTTTC 3420
Db      3361 AATAGTATGAACTTTTCTATCTATTTAGTGATATATTTTATGATGGAATAAACTTTC 3420
Qy      3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3453
Db      3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3453

```

## RESULT 8

AAQ79068

ID AAQ79068

XX AAQ79068 standard; cDNA; 3453 BP.

AC AAQ79068;

XX

DT 25-MAR-2003 (updated)

DT 04-JUL-1995 (first entry)

XX

DE Mouse flk-2 cDNA.

XX

KW Fetal liver kinase-2; flk-2; protein tyrosine-kinase receptor;

KW hematopoiesis; stem cell; ds.

XX

OS Mus sp.

XX

FH Key Location/Qualifiers

FT CDS 58..3039

XX

XX

XX

XX

XX

XX

```

FT      sig_peptide /*tag= a
FT      mat_peptide /*tag= b
FT      /*tag= c
XX      US5367057-A.
XX      22-NOV-1994.
PD      30-APR-1993; 93US-0055269.
XX      02-APR-1991; 91US-0679666.
PR      28-JUN-1991; 91US-0728913.
PR      15-NOV-1991; 91US-0793065.
PR      24-DEC-1991; 91US-0813593.
PR      26-JUN-1992; 92US-0906397.
PR      12-NOV-1992; 92US-0975049.
PR      19-NOV-1992; 92US-0977451.
PR      30-APR-1993; 93US-0055269.
XX      (UYPR-) UNIV PRINCETON.
PA      Lemischka IR;
XX      WPI; 1995-005894/01.
DR      P-PSDB; AAR67535.
XX      Murine flk-2 receptor protein tyrosine kinase - used to stimulate
PT      proliferation and/or stimulation of primitive mammalian
PT      hematopoietic stem cells in vitro or in vivo.
XX      Disclosure; Fig. 1A-1F; 69pp; English.
XX      cDNAs encoding receptor protein tyrosine-kinases, mouse fetal liver
CC      kinase-2 (flk-2), human flk-2 and mouse flk-1 are given in AAQ79068-70,
CC      respectively, and the deduced amino acid sequences in AAR67535-37,
CC      respectively.
CC      (Updated on 25-MAR-2003 to correct PF field.)
XX      Sequence 3453 BP; 946 A; 822 C; 876 G; 809 T; 0 other;
SQ

```

## Query Match

Best Local Similarity 100.0%; Score 3451.4; DB 16; Length 3453;

Matches 3452; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCGGCTGGCTACCGCGCTCCGGAGCCATCGGGCGTTGGCGCAGCGAGCCGG 60

Db 1 GCGGCTGGCTACCGCGCTCCGGAGCCATCGGGCGTTGGCGCAGCGAGCCGG 60

Qy 61 CCGCTGCTGCTGCTGTTGTTTCTCAGTAATGATTCTTGAGACGGTTACAAACAGAC 120

Db 61 CCGCTGCTGCTGCTGTTGTTTCTCAGTAATGATTCTTGAGACGGTTACAAACAGAC 120

Qy 121 CTGCTGTGATCAAGTGTGTTTAACTCAGTCAAGAACTGCTCATCAGCGGGAAG 180

Db 121 CTGCTGTGATCAAGTGTGTTTAACTCAGTCAAGAACTGCTCATCAGCGGGAAG 180

Qy 181 CCATCATCTGATCCGAATGTTGCGAGGATCCCAAGAACCTCCAGTGTACCCGAGCGC 240

Db 181 CCATCATCTGATCCGAATGTTGCGAGGATCCCAAGAACCTCCAGTGTACCCGAGCGC 240

Qy 241 CAGAGTGAAGGAGCGGTATATGAAGCGGCCACCGTGGAGTGGCGGATCTGGGTCCATC 300

Db 241 CAGAGTGAAGGAGCGGTATATGAAGCGGCCACCGTGGAGTGGCGGATCTGGGTCCATC 300

Qy 301 ACCCTGCAAGTGCAGCTCGCCACCCAGGGGACCTTTCTGCTCTGGGTCTTTAAGCAC 360

Db 301 ACCCTGCAAGTGCAGCTCGCCACCCAGGGGACCTTTCTGCTCTGGGTCTTTAAGCAC 360

Qy 361 AGCTCCCTGGGTGCGAGCGGCACCTTTGATTTTACAAAACAGAGGAATCGTTTCCATGGCC 420

Db 361 AGCTCCCTGGGTGCGAGCGGCACCTTTGATTTTACAAAACAGAGGAATCGTTTCCATGGCC 420

QY 421 ATCTTGAAGCTGACAGAGACCCAGGACGAGAGATACCTACTCCATATTCAGAGCGAAGCG 480  
DB |||||  
QY 421 ATCTTGAAGCTGACAGAGACCCAGGACGAGAGATACCTACTCCATATTCAGAGCGAAGCG 480  
DB |||||  
QY 481 GCCAACTACACAGTACTGTTTCACAGTGAATGTAAGAGATACACAGCTGTATGCTTAAGG 540  
DB |||||  
QY 481 GCCAACTACACAGTACTGTTTCACAGTGAATGTAAGAGATACACAGCTGTATGCTTAAGG 540  
DB |||||  
QY 541 AGACCTTACTTTAGGAAGATGAAAAACAGGATGCACTGCTGTGCATCTCCGAGGGTGT 600  
DB |||||  
QY 541 AGACCTTACTTTAGGAAGATGAAAAACAGGATGCACTGCTGTGCATCTCCGAGGGTGT 600  
DB |||||  
QY 601 CCGGAGCCCACTGTGGAGTGGGTCTCTGACGCTCCACAGGGAAGCTGTAAGNAGAA 660  
DB |||||  
QY 601 CCGGAGCCCACTGTGGAGTGGGTCTCTGACGCTCCACAGGGAAGCTGTAAGNAGAA 660  
DB |||||  
QY 661 GGCCTCTGTTGTTCAGAAAGAGGAAAAAGGTACTTTCATGAGTGTTCGGAACAGACATC 720  
DB |||||  
QY 661 GGCCTCTGTTGTTCAGAAAGAGGAAAAAGGTACTTTCATGAGTGTTCGGAACAGACATC 720  
DB |||||  
QY 721 AGATGCTGTGCTAGAAAATGCACTGGGCCGCGAATGCAACCAAGCTGTTCACATAGATCTA 780  
DB |||||  
QY 721 AGATGCTGTGCTAGAAAATGCACTGGGCCGCGAATGCAACCAAGCTGTTCACATAGATCTA 780  
DB |||||  
QY 781 AACCAGGCTCTCAGAGCACACTGCCCCAGTTATTCCTGAAAGTGGGGAAACCTTGTGG 840  
DB |||||  
QY 781 AACCAGGCTCTCAGAGCACACTGCCCCAGTTATTCCTGAAAGTGGGGAAACCTTGTGG 840  
DB |||||  
QY 841 ATCAGGTGTAAGGCCCATCATGTGAAACCATGGAATTCGGGCTCACCTGGGAGCTGGAAGAC 900  
DB |||||  
QY 841 ATCAGGTGTAAGGCCCATCATGTGAAACCATGGAATTCGGGCTCACCTGGGAGCTGGAAGAC 900  
DB |||||  
QY 901 AAAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATG 960  
DB |||||  
QY 901 AAAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATG 960  
DB |||||  
QY 961 ATTCGGATTTCTCTGSCCTTTGTGTCTTCGTTGGGAAGGAAACGACACCGGATTAACACC 1020  
DB |||||  
QY 961 ATTCGGATTTCTCTGSCCTTTGTGTCTTCGTTGGGAAGGAAACGACACCGGATTAACACC 1020  
DB |||||  
QY 1021 TGCTCTTCTCCTCAAGACCCAGCCAGTCAGCTGCTGAGCCATCTCTAGAAAAGGGTTT 1080  
DB |||||  
QY 1021 TGCTCTTCTCCTCAAGACCCAGCCAGTCAGCTGCTGAGCCATCTCTAGAAAAGGGTTT 1080  
DB |||||  
QY 1081 ATAAAACGCTACCAAGCTCGCAAGAAAGATGAAATTTGACCCGTAACGAAAGTTCTGCTTC 1140  
DB |||||  
QY 1081 ATAAAACGCTACCAAGCTCGCAAGAAAGATGAAATTTGACCCGTAACGAAAGTTCTGCTTC 1140  
DB |||||  
QY 1141 TCAGTCAGGTTTAAAGCGTACCCACGAATCCGATGCACTGGATCTCTCTCAAGCCTCA 1200  
DB |||||  
QY 1141 TCAGTCAGGTTTAAAGCGTACCCACGAATCCGATGCACTGGATCTCTCTCAAGCCTCA 1200  
DB |||||  
QY 1201 TTTCTTGTGAAACAGAGAGGCTGGAGGATGGGTACAGCATATCTAAATTTTGGCATAT 1260  
DB |||||  
QY 1201 TTTCTTGTGAAACAGAGAGGCTGGAGGATGGGTACAGCATATCTAAATTTTGGCATAT 1260  
DB |||||  
QY 1261 AAGAAACAGCCAGGAGGATGATATTTCTATGCAAGAAATGATGACGCCCGAGTTTCAACAAA 1320  
DB |||||  
QY 1261 AAGAAACAGCCAGGAGGATGATATTTCTATGCAAGAAATGATGACGCCCGAGTTTCAACAAA 1320  
DB |||||  
QY 1321 ATGTTTCAAGCTGAATAAAGAAAGAAACCTCAAGTGTCTAGCAAAATGCTCAGCCAGCCAG 1380  
DB |||||  
QY 1381 GCGTCTGTTCTCTGATGGCTACCCGCTACCCCTCTTGACCTGGGAAGAGTTCGGAC 1440  
DB |||||  
QY 1381 GCGTCTGTTCTCTGATGGCTACCCGCTACCCCTCTTGACCTGGGAAGAGTTCGGAC 1440  
DB |||||  
QY 1441 AAATCTCCAAATGACGAGGAAATCCAGAGGAGTTTGGAAATAAAAGGCTAACAGA 1500  
DB |||||  
QY 1441 AAATCTCCAAATGACGAGGAAATCCAGAGGAGTTTGGAAATAAAAGGCTAACAGA 1500  
DB |||||

QY 1501 AAAAGTGTTCGCGAGTGGGTGTCGAGCAGTACTCTAAAATATGAGTGAAGCCGCGGAAAGGG 1560  
DB |||||  
QY 1501 AAAAGTGTTCGCGAGTGGGTGTCGAGCAGTACTCTAAAATATGAGTGAAGCCGCGGAAAGGG 1560  
DB |||||  
QY 1561 CTTCTGGTCAAAATGCTGTGCGTACAAATTTCTATGCGGCACTCTTTCGGAACCAATCTTTTAA 1620  
DB |||||  
QY 1561 CTTCTGGTCAAAATGCTGTGCGTACAAATTTCTATGCGGCACTCTTTCGGAACCAATCTTTTAA 1620  
DB |||||  
QY 1621 AACTCACAGGCCCCCTTCCCTTTTCATCCNAGAACATCTCTCTTCTATGCGACCATTTGG 1680  
DB |||||  
QY 1621 AACTCACAGGCCCCCTTCCCTTTTCATCCNAGAACATCTCTCTTCTATGCGACCATTTGG 1680  
DB |||||  
QY 1681 CTTCTGCTCCCCCTTCAATTTGTTCTCATTTGTTGATCTGCCCCAACAAATACAAAAAGCAA 1740  
DB |||||  
QY 1681 CTTCTGCTCCCCCTTCAATTTGTTCTCATTTGTTGATCTGCCAACAAATACAAAAAGCAA 1740  
DB |||||  
QY 1741 TTTAGTACAGAGTCAAGTGCAGATGATCCAGGTGATCCAGGTGAGTTCGCGAGAGAGAAC 1800  
DB |||||  
QY 1741 TTTAGTACAGAGTCAAGTGCAGATGATCCAGGTGATCCAGGTGAGTTCGCGAGAGAGAAC 1800  
DB |||||  
QY 1801 TTTAGTGTGACTTTCAGGAGCTATGAAATATGACTTTAAGTGGAGTTCGCGAGAGAGAAC 1860  
DB |||||  
QY 1801 TTTAGTGTGACTTTCAGGAGCTATGAAATATGACTTTAAGTGGAGTTCGCGAGAGAGAAC 1860  
DB |||||  
QY 1861 TTTAGAGTGTTCGGAAGGCTCTGCGGCTCTGCGGAGGCTGATGAACGCCACCGGCC 1920  
DB |||||  
QY 1861 TTTAGAGTGTTCGGAAGGCTCTGCGGCTCTGCGGAGGCTGATGAACGCCACCGGCC 1920  
DB |||||  
QY 1921 TATGCACTTAGTAAAAACGGGAGTCTCAATTCAGGTGCGCGTGAAGATGCTTAAAGAGAAA 1980  
DB |||||  
QY 1921 TATGCACTTAGTAAAAACGGGAGTCTCAATTCAGGTGCGCGTGAAGATGCTTAAAGAGAAA 1980  
DB |||||  
QY 1981 GCTGACAGCTGTGAAAAAGAAAGCTCTCATGTGCGAGCTCAAAATATGATGACCCCTGGGA 2040  
DB |||||  
QY 1981 GCTGACAGCTGTGAAAAAGAAAGCTCTCATGTGCGAGCTCAAAATATGATGACCCCTGGGA 2040  
DB |||||  
QY 2041 CACCATGCAACATCGTGAATCTGCTGGGGGATGCACACTGTTCAGGGGCGAGTGTACTTG 2100  
DB |||||  
QY 2041 CACCATGCAACATCGTGAATCTGCTGGGGGATGCACACTGTTCAGGGGCGAGTGTACTTG 2100  
DB |||||  
QY 2101 ATTTTGTGAATTTGTTGCTATGCTGACCTCTCACTACCTAAGAGTAAAAAGAGAGAAG 2160  
DB |||||  
QY 2101 ATTTTGTGAATTTGTTGCTATGCTGACCTCTCACTACCTAAGAGTAAAAAGAGAGAAG 2160  
DB |||||  
QY 2161 TTTTCAAGSACATGACAGAGATTTTAAAGSACATTAATTTTCAAGTCTTACCTACTTTTC 2220  
DB |||||  
QY 2161 TTTTCAAGSACATGACAGAGATTTTAAAGSACATAATTTTCAAGTCTTACCTACTTTTC 2220  
DB |||||  
QY 2221 CAGGACATTTCAAAATTCAGCATGCTGTTTCAAGAGAGTTCAGTTACACCCGCGCTTG 2280  
DB |||||  
QY 2221 CAGGACATTTCAAAATTCAGCATGCTGTTTCAAGAGAGTTCAGTTACACCCGCGCTTG 2280  
DB |||||  
QY 2281 GATCAGCTCTCAGGGTTCAATGCGGAATTCATTTCTGAGATGAGATGGAATATGAA 2340  
DB |||||  
QY 2281 GATCAGCTCTCAGGGTTCAATGCGGAATTCATTTCTGAGATGAGATGGAATATGAA 2340  
DB |||||  
QY 2341 AACCAAGAGAGGCTGCGAGAGAGAGAGAGAGATTTTGAACGCTGCTGAGCTTTGAAGAC 2400  
DB |||||  
QY 2341 AACCAAGAGAGGCTGCGAGAGAGAGAGAGAGATTTTGAACGCTGCTGAGCTTTGAAGAC 2400  
DB |||||  
QY 2401 CTTCTTGTCTTTGCGTACCAAGTGGCAAAAGGATGGAATTTCTGAGTTCAGTCTGCTGT 2460  
DB |||||  
QY 2401 CTTCTTGTCTTTGCGTACCAAGTGGCAAAAGGATGGAATTTCTGAGTTCAGTCTGCTGT 2460  
DB |||||  
QY 2461 GTCCACAGAGACTGCGACGAGGATGTTGTTGCTCACCCAGGGAAGTGTGGAAGATC 2520  
DB |||||  
QY 2461 GTCCACAGAGACTGCGACGAGGATGTTGTTGCTCACCCAGGGAAGTGTGGAAGATC 2520  
DB |||||  
QY 2521 TGTGACTTTTGGACTGGCCGAGACATCTCAGGAGTCCAGCTACCTCAGGAGGCAAC 2580  
DB |||||  
QY 2521 TGTGACTTTTGGACTGGCCGAGACATCTCAGGAGTCCAGCTACCTCAGGAGGCAAC 2580  
DB |||||  
QY 2581 GCACGGCTCGCGGTGAAGTGGATGGCAACCGAGAGCTTATTTGAAGGGATCTTACAAATC 2640  
DB |||||

```

Db 2581 GCACGGCTGCCGGTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGATCTACACATC 2640
Qy 2641 AAGAGTCAGCTCTGCTCTCGGATCTCTCTCTGGAGATATTTTCACTGGGTGTGAAC 2700
Db 2641 AAGAGTCAGCTCTGCTCTCGGATCTCTCTCTGGAGATATTTTCACTGGGTGTGAAC 2700
Qy 2701 CCTTACCTGGCATTTCTCTCGAGCTTAACTTTCTATAAATGATTCAGAGTGGATTTAAA 2760
Db 2701 CCTTACCTGGCATTTCTCTCGAGCTTAACTTTCTATAAATGATTCAGAGTGGATTTAAA 2760
Qy 2761 ATGAGCAGCATTTCTATGCGACAGAGGATATCTTTGTAATGCAATCTCTGCTGGCT 2820
Db 2761 ATGAGCAGCATTTCTATGCGACAGAGGATATCTTTGTAATGCAATCTCTGCTGGCT 2820
Qy 2821 TTTGACTCAAGGAGCGGCTCTCTCTCCCACTGACTTCAATTTTAGGATGTACGCTG 2880
Db 2821 TTTGACTCAAGGAGCGGCTCTCTCTCCCACTGACTTCAATTTTAGGATGTACGCTG 2880
Qy 2881 GCAGAGGCGAGAGCATGTATCAGAAATCATCCATCTACCAAAACAGGCGGCCCT 2940
Db 2881 GCAGAGGCGAGAGCATGTATCAGAAATCATCCATCTACCAAAACAGGCGGCCCT 2940
Qy 2941 CAGCAGAGGCGGCTCAGAGCCAGTGCACAGCCAGGTCAGAGTTCACAGAGAA 3000
Db 2941 CAGCAGAGGCGGCTCAGAGCCAGTGCACAGCCAGGTCAGAGTTCACAGAGAA 3000
Qy 3001 AGAAGTTAGCGAGGCGCTTGACCCCGCCAGCTAGCAGGCTGTAGACCGCAGAGCCA 3060
Db 3001 AGAAGTTAGCGAGGCGCTTGACCCCGCCAGCTAGCAGGCTGTAGACCGCAGAGCCA 3060
Qy 3061 AGATTAGCTCGCTCTGAGGAGCGCCCTACAGCGGCTTCTGCTGACTTTTCTCT 3120
Db 3061 AGATTAGCTCGCTCTGAGGAGCGCCCTACAGCGGCTTCTGCTGACTTTTCTCT 3120
Qy 3121 AGATGCTGTGCCATTACTCCAAAGTGACTTCTATAAATCAAACTCTCTCGCACAG 3180
Db 3121 AGATGCTGTGCCATTACTCCAAAGTGACTTCTATAAATCAAACTCTCTCGCACAG 3180
Qy 3181 GCGGAGGAGCCCAATAGAGACTTGTGTGAGCGCCCTACCTGCGGCGCTTTCCAGC 3240
Db 3181 GCGGAGGAGCCCAATAGAGACTTGTGTGAGCGCCCTACCTGCGGCGCTTTCCAGC 3240
Qy 3241 AGCTTGAGGAGGAGCCATGTATCTGAATATAGTATATCTGTGAATACGTGAAACAA 3300
Db 3241 AGCTTGAGGAGGAGCCATGTATCTGAATATAGTATATCTGTGAATACGTGAAACAA 3300
Qy 3301 ACCAAACCGTTTTTGTCTAAGGAGAGCTAAATATGATTTTAAATAATCTATGTTTAA 3360
Db 3301 ACCAAACCGTTTTTGTCTAAGGAGAGCTAAATATGATTTTAAATAATCTATGTTTAA 3360
Qy 3361 AATACTATGTAACCTTTTCACTATTTAGTGATATATTTATGATGGAATAAATCTTC 3420
Db 3361 AATACTATGTAACCTTTTCACTATTTAGTGATATATTTATGATGGAATAAATCTTC 3420
Qy 3421 TACTGTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3453
Db 3421 TACTGTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3453

```

## RESULT 9

AAQ29954

ID AAQ29954 standard; cDNA; 3453 BP.

XX XX

AC AAQ29954;

XX XX

DT 25-MAR-2003 (updated)

DT 15-MAR-1993 (first entry)

XX XX

DE Murine flk-2 cDNA sequence.

XX XX

KW Thymidine kinase; TK; haematopoietic; stem cells; proliferation;

KW differentiation; progenitor cells; foetal liver kinase; ss.

..

```

XX OS Mus musculus.
XX FH Key
XX FT CDS
XX FT Location/Qualifiers
XX FT 31..3009
XX FT /*tag= a
XX PN W09217486-A1.
XX PD 15-OCT-1992.
XX PF 02-APR-1992; 92WO-US02750.
XX PR 02-APR-1991; 91US-0679666.
XX PR 28-JUN-1991; 91US-0728913.
XX PR 15-NOV-1991; 91US-0793065.
XX PR 24-DEC-1991; 91US-0813593.
XX PA (UYPR-) UNIV PRINCETON.
XX PI Lemischka IR;
XX PI WPI; 1992-366185/44.
XX DR P-PSDB; AAR28038.
XX ST Stimulating proliferation and/or differentiation of primitive
XX PT mammalian haematopoietic stem cells - using ligand that binds
XX PT thymidine kinase and flk-1 and flk-2
XX PS Claim 6; Fig 1a; 94pp; English.
XX CC The murine foetal liver kinase (flk) -2 clone was isolated by
XX CC standard PCR techniques from stem-cell receptor-contg. tissue cDNA
XX CC libraries. Suitable tissues include foetal liver, spleen or thymus
XX CC cells or adult marrow or brain cells. The PCR primers used are based
XX CC on known sections of the flk-2 gene. The murine flk-2 clone may be
XX CC used in a vector to transform haematopoietic cells. The thymidine
XX CC kinase encoded by flk-2 is expressed in primitive but not mature
XX CC haematopoietic cells. Ligand binding to the TK may be prepd. which
XX CC can stimulate proliferation and/or differentiation of primitive
XX CC haematopoietic cells in vivo. The ligands can stimulate the
XX CC proliferation of additional primitive stem cells, differentiation into
XX CC more mature progenitor cells, or both.
XX CC See also AAQ29955-7.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 3453 BP; 946 A; 822 C; 876 G; 809 T; 0 other;

Query Match 99.9%; Score 3448.2; DB 13; Length 3453;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3450; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GCGGCTGGCTACCGCGCTCCGAGGCCATGCGGGCTTGGCGCAGCGCAGCGCGG 60
Db 1 GCGGCTGGCTACCGCGCTCCGAGGCCATGCGGGCTTGGCGCAGCGCAGCGCGG 60
Qy 61 CGGCTGCTGCTGCTGTTGTTTCTCAGTAATGATTTTGAGACCGTTACAAACCAAGAC 120
Db 61 CGGCTGCTGCTGCTGTTGTTTCTCAGTAATGATTTTGAGACCGTTACAAACCAAGAC 120
Qy 121 CTGCTGTGATCAAGTGTGTTTAAATCATCATCAGAACATGCTCATCAGCGGGAAG 180
Db 121 CTGCTGTGATCAAGTGTGTTTAAATCATCATCAGAACATGCTCATCAGCGGGAAG 180
Qy 181 CCATCATGTACCGAATGGTGGAGGATCCCAAGACCTTCCAGTGTACCCCGAGGCGC 240
Db 181 CCATCATGTACCGAATGGTGGAGGATCCCAAGACCTTCCAGTGTACCCCGAGGCGC 240
Qy 241 CAGAGTGAAGGAGCGGTATATGAGCGGCCACCGTGGAGGTGGCGGATCTGGGTCCATC 300
Db 241 CAGAGTGAAGGAGCGGTATATGAGCGGCCACCGTGGAGGTGGCGGATCTGGGTCCATC 300
Qy 301 ACCGTGCAAGTGCAGCTCGCCACCCAGGGGAGCCTTTCTCTGCTCTGTTAAAGCAC 360

```

|||||  
Db 301 ACCCTGCAAGTGCAGCTCGCCACCCGAGGGACCTTTCCTGCCCTCGGGTCTTTAAGCAC 360  
Qy 361 AGCTCCCTGGGCTGCCGCCGCACTTTGATTTACAAAACAGAGAAATCGTTTCCATGGCC 420  
Db 361 AGCTCCCTGGGCTGCCGCCGCACTTTGATTTACAAAACAGAGAAATCGTTTCCATGGCC 420  
Qy 421 ATCTTGAACTGACAGAGACCCAGGCGAGGAGAAATACCTACTCCATATTCAGAGCGAACGC 480  
Db 421 ATCTTGAACTGACAGAGACCCAGGCGAGGAGAAATACCTACTCCATATTCAGAGCGAACGC 480  
Qy 481 GCCAACTACAGTACTGTTTCCACAGTGAATGTAAGAGATACACAGCTGTATGTGCTAAGG 540  
Db 481 GCCAACTACAGTACTGTTTCCACAGTGAATGTAAGAGATACACAGCTGTATGTGCTAAGG 540  
Qy 541 AGACCTTTACTTTAGGAAGATGGAAGAACCCAGGATGCTCTCTGCATCTCCGAGGGTGT 600  
Db 541 AGACCTTTACTTTAGGAAGATGGAAGAACCCAGGATGCTCTCTGCATCTCCGAGGGTGT 600  
Qy 601 CCGAGGCCACTGTGGAGTGGTGTCTGACGCTCCGACGGGAAAGCTGTAAAGAGAA 660  
Db 601 CCGAGGCCACTGTGGAGTGGTGTCTGACGCTCCGACGGGAAAGCTGTAAAGAGAA 660  
Qy 661 GGCCTGCTGTGTCAGAAAGGAGAAAGGTACTTCAATGAGTGTTCGGGACAGACATC 720  
Db 661 GGCCTGCTGTGTCAGAAAGGAGAAAGGTACTTCAATGAGTGTTCGGGACAGACATC 720  
Qy 721 AGATGCTGTGTAGAAATGCACTGGGCGCGAATGCAACCAAGCTGTCCACATAGATCTA 780  
Db 721 AGATGCTGTGTAGAAATGCACTGGGCGCGAATGCAACCAAGCTGTCCACATAGATCTA 780  
Qy 781 AACAGGCTCTCTAGAGACACTGCCCCAGTTATTCCTGAAAGTGGGGGAAACCTTTGTGG 840  
Db 781 AACAGGCTCTCTAGAGACACTGCCCCAGTTATTCCTGAAAGTGGGGGAAACCTTTGTGG 840  
Qy 841 ATCAGGTGTAAAGCCATCCATGTGAACCATGGAATTCGGGCTCACTGGGAGCTGGAGAC 900  
Db 841 ATCAGGTGTAAAGCCATCCATGTGAACCATGGAATTCGGGCTCACTGGGAGCTGGAGAC 900  
Qy 901 AAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAGTACTACTCCACAAACAGAGACCATG 960  
Db 901 AAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAGTACTACTCCACAAACAGAGACCATG 960  
Qy 961 ATTCGGATTTCTTTGGCCTTTGTGCTTCCGTTGGGAGGAAACGACACCGGATATTACACC 1020  
Db 961 ATTCGGATTTCTTTGGCCTTTGTGCTTCCGTTGGGAGGAAACGACACCGGATATTACACC 1020  
Qy 1021 TGCTCTTCTCAAAGCACCCAGCCAGTCAGGTTGGTGAACATCCTAGAAAAAGGGTTT 1080  
Db 1021 TGCTCTTCTCAAAGCACCCAGCCAGTCAGGTTGGTGAACATCCTAGAAAAAGGGTTT 1080  
Qy 1081 ATAAACGCTACCGCTCGCAAGAGAGATATGAATTTGACCCGTACGAAAGTTCTGCTTC 1140  
Db 1081 ATAAACGCTACCGCTCGCAAGAGAGATATGAATTTGACCCGTACGAAAGTTCTGCTTC 1140  
Qy 1141 TCAGTCAGGTTTAAAGCGTACCCACGATCCGATGCACTGGATCTTCTCAAGCCTCA 1200  
Db 1141 TCAGTCAGGTTTAAAGCGTACCCACGATCCGATGCACTGGATCTTCTCAAGCCTCA 1200  
Qy 1201 TTTCTTTGTGAACAGAGAGGCTTGAGAGTGGGTACAGCATATCTAAATTTTGGCATCAT 1260  
Db 1201 TTTCTTTGTGAACAGAGAGGCTTGAGAGTGGGTACAGCATATCTAAATTTTGGCATCAT 1260  
Qy 1261 AAGAACAGCCAGGAGGATCATATTTCTATCGAAGAAATGATGCGCCAGTTCCACCAA 1320  
Db 1261 AAGAACAGCCAGGAGGATCATATTTCTATCGAAGAAATGATGAGCCAGTTCCACCAA 1320  
Qy 1321 ATGTTTCAGCTGAATATGAAGAAAGAAACCTCAAGTGTAGCAAAATGCTCAGCCAGCCAG 1380  
Db 1321 ATGTTTCAGCTGAATATGAAGAAAGAAACCTCAAGTGTAGCAAAATGCTCAGCCAGCCAG 1380  
Qy 1381 GCGTCTGTCTCTGATGGCTACCCGCTACCCCTCTTGGAACCTTGGAAGAGTGTTCGGAC 1440  
Db 1381 GCGTCTGTCTCTGATGGCTACCCGCTACCCCTCTTGGAAGAGTGTTCGGAC 1440

1381 GCGTCTGTCTCTGATGGCTACCCGCTACCCCTTTGGAGCCTGGAAGAGTGTTCGGAC 1440  
Qy 1441 AAATCTCCCAATTCACGAGGAAATCCACAGAGAGTTCGGAATAAAGAGCTTAACAGA 1500  
Db 1441 AAATCTCCCAATTCACGAGGAAATCCACAGAGAGTTCGGAATAAAGAGCTTAACAGA 1500  
Qy 1501 AAAGTGTTCGCCAGTGGTGTGAGCAGTACTCTAAATATGAGTGAGGCGGGGAAAGGG 1560  
Db 1501 AAAGTGTTCGCCAGTGGTGTGAGCAGTACTCTAAATATGAGTGAGGCGGGGAAAGGG 1560  
Qy 1561 CTTCTGGTCAAAATCTGTGCGTACAAATTCATATGGGCAAGTCTTTCGAAAACCATCTTTTA 1620  
Db 1561 CTTCTGGTCAAAATCTGTGCGTACAAATTCATATGGGCAAGTCTTTCGAAAACCATCTTTTA 1620  
Qy 1621 AACTCACAGGCCCCCTTCCCTTTTCATCCAAACAACATCTCTCTATGCGACCAATGGG 1680  
Db 1621 AACTCACAGGCCCCCTTCCCTTTTCATCCAAACAACATCTCTCTATGCGACCAATGGG 1680  
Qy 1681 CTCTGTCTCCCTTCATTGTTTCTCAATCTGTGTGATCTGCGCAAAATACAAAAGCAA 1740  
Db 1681 CTCTGTCTCCCTTCATTGTTTCTCAATCTGTGTGATCTGCGCAAAATACAAAAGCAA 1740  
Qy 1741 TTTAGGTACAGAGTCACTGCAGATGATCCAGGTGAATGGCCCCCTGGATAACAGTAC 1800  
Db 1741 TTTAGGTACAGAGTCACTGCAGATGATCCAGGTGAATGGCCCCCTGGATAACAGTAC 1800  
Qy 1801 TTTACGTTGATCTTCAGGCACTATGAATATGACCTTAAAGTGGGAGTTCGAGAGAGAAC 1860  
Db 1801 TTTACGTTGATCTTCAGGCACTATGAATATGACCTTAAAGTGGGAGTTCGAGAGAGAAC 1860  
Qy 1861 TTAGAGTTTGGGAGGTCCTGGGCTGCGGCTTTGGGAGGGTGAATGAACCCACGGCC 1920  
Db 1861 TTAGAGTTTGGGAGGTCCTGGGCTGCGGCTTTGGGAGGGTGAATGAACCCACGGCC 1920  
Qy 1921 TATGGCATTTAGTAAACCGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAGAGAAA 1980  
Db 1921 TATGGCATTTAGTAAACCGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAGAGAAA 1980  
Qy 1981 GCTCAGAGCTGTGAAAAAGAGCTCTCATGTGGAGCTCAAAATGATGATGATGATGATG 2040  
Db 1981 GCTCAGAGCTGTGAAAAAGAGCTCTCATGTGGAGCTCAAAATGATGATGATGATGATG 2040  
Qy 2041 CACATGACACATCTGTGAATCTGCTGGGGCATGACACATGTCAGGGCCAGTGTACTTG 2100  
Db 2041 CACATGACACATCTGTGAATCTGCTGGGGCATGACACATGTCAGGGCCAGTGTACTTG 2100  
Qy 2101 ATTTTGAATATGTTGCTATGGTGACCTCTCAACTACCTAAAGAAATGAAAGAGAGAG 2160  
Db 2101 ATTTTGAATATGTTGCTATGGTGACCTCTCAACTACCTAAAGAAATGAAAGAGAGAG 2160  
Qy 2161 TTTTCAAGGACATGAGAGATTTTAAAGGAACATAATTTTCACTTCTTACTTCTTCTTC 2220  
Db 2161 TTTTCAAGGACATGAGAGATTTTAAAGGAACATAATTTTCACTTCTTACTTCTTCTTC 2220  
Qy 2221 CAGGACATTTCAAATTCAGCATGCTGTTTCAAGAGTTCAGTTTACACCCGCTTTC 2280  
Db 2221 CAGGACATTTCAAATTCAGCATGCTGTTTCAAGAGTTCAGTTTACACCCGCTTTC 2280  
Qy 2281 GATCAGCTCTCAGGTTCAATGGGAAATTCATTTCAATCTGAGATGAGATTTGAATATGAA 2340  
Db 2281 GATCAGCTCTCAGGTTCAATGGGAAATTCATTTCAATCTGAGATGAGATTTGAATATGAA 2340  
Qy 2341 AACCAAGAGGCTGGCAG 2400  
Db 2341 AACCAAGAGGCTGGCAG 2400  
Qy 2401 CTCCTTTGCTTTCAGTACCAAGTGGCCAAAGGCAATGGAATTCCTGGAGTTCAAGTCTGT 2460  
Db 2401 CTCCTTTGCTTTCAGTACCAAGTGGCCAAAGGCAATGGAATTCCTGGAGTTCAAGTCTGT 2460  
Qy 2461 GTCCACAGAGACCTGGCAG 2520  
Db 2461 GTCCACAGAGACCTGGCAG 2520





121 CTGCGTGTGATCAAGTGTGTTTTAAATCAGTCAATGAGAACAAATGCTCATCAGCGGGAAG 180  
121 CTGCGTGTGATCAAGTGTGTTTTAAATCAGTCAATGAGAACAAATGCTCATCAGCGGGAAG 180  
181 CCATCATCGTACCGAATGGTGGAGGATCCCGAAGACCTCCAGTGTACCCCGAGCGC 240  
181 CCATCATCGTACCGAATGGTGGAGGATCCCGAAGACCTCCAGTGTACCCCGAGCGC 240  
241 CAGAGTGAAGGAGCGGTATATGAAGCGGCGACCGTGGAGGTGGCGAGTCTGGTTCATC 300  
241 CAGAGTGAAGGAGCGGTATATGAAGCGGCGACCGTGGAGGTGGCGAGTCTGGTTCATC 300  
301 ACCCTGCAAGTGCAGCTCGCCACCCAGGAGGACCTTTCTGCTCTGGTCTTTAAGCAC 360  
301 ACCCTGCAAGTGCAGCTCGCCACCCAGGAGGACCTTTCTGCTCTGGTCTTTAAGCAC 360  
361 AGCTCCCTGGGCTGCCAGCGCACTTTGATTTACAAACAGAGAAATCGTTTCCATGGCC 420  
361 AGCTCCCTGGGCTGCCAGCGCACTTTGATTTACAAACAGAGAAATCGTTTCCATGGCC 420  
421 ATCTTGAACTGACAGAGACCCAGGAGGAGATACCTACTCCATATTCAGAGCGAAGC 480  
421 ATCTTGAACTGACAGAGACCCAGGAGGAGATACCTACTCCATATTCAGAGCGAAGC 480  
481 GCCAACTACACAGTACTGTTTACAGTGAATGTAAGAGATACACAGCTGTATGCTTAAG 540  
481 GCCAACTACACAGTACTGTTTACAGTGAATGTAAGAGATACACAGCTGTATGCTTAAG 540  
541 AGACCTTACTTTAGGAAGATGGAACCCAGAGTCACTGCTCTGCATCTCCGAGGGTGT 600  
541 AGACCTTACTTTAGGAAGATGGAACCCAGAGTCACTGCTCTGCATCTCCGAGGGTGT 600  
601 CCGGAGCCCACTGTGGAGTGGTCTCTGAGTCCCAAGTCCCAAGGAAAGCTGTAAAGAGAA 660  
601 CCGGAGCCCACTGTGGAGTGGTCTCTGAGTCCCAAGTCCCAAGGAAAGCTGTAAAGAGAA 660  
661 GGCCCTGCTGTGTCAGAAAGGAGGAAAGTACTTATGAGTGTTCGGAACAGACATC 720  
661 GGCCCTGCTGTGTCAGAAAGGAGGAAAGTACTTATGAGTGTTCGGAACAGACATC 720  
721 AGATGCTGTGTCAGAAATGCACTGGGCGGAAATGCAACCAAGTGTTCACCATAGATCTA 780  
721 AGATGCTGTGTCAGAAATGCACTGGGCGGAAATGCAACCAAGTGTTCACCATAGATCTA 780  
781 AACAGGCTCTCTAGAGACACACTGCCCCAGTATTTCTGAAAGTGGGGAAACCTTTGG 840  
781 AACAGGCTCTCTAGAGACACACTGCCCCAGTATTTCTGAAAGTGGGGAAACCTTTGG 840  
841 ATCAGGTGAAGGCCATCCATGTGAACCATGGAATTCGGGCTCACTGGAGCTGGAAGAC 900  
841 ATCAGGTGAAGGCCATCCATGTGAACCATGGAATTCGGGCTCACTGGAGCTGGAAGAC 900  
901 AAAGCCCTGGAGGAGGAGCTACTTTGAGATGAGTACTTCTCCACAAACAGGACCATG 960  
901 AAAGCCCTGGAGGAGGAGCTACTTTGAGATGAGTACTTCTCCACAAACAGGACCATG 960  
961 ATTCCGATTTCTTGGCCCTTTGCTCTTCCGTTGGGAAGGAAACGACACCGGATATTAC 1020  
961 ATTCCGATTTCTTGGCCCTTTGCTCTTCCGTTGGGAAGGAAACGACACCGGATATTAC 1020  
1021 TGCTCTTCCCTCAAGACACCCAGCGAGTCAAGCTGAGTGTGGAACATCTAGAAAGGGTTT 1080  
1021 TGCTCTTCCCTCAAGACACCCAGCGAGTCAAGCTGAGTGTGGAACATCTAGAAAGGGTTT 1080  
1081 ATAAACGCTACAGCTCCGAGAGAGTATGAATTTGACCTGAGGAAAGTTCTGCTTC 1140  
1081 ATAAACGCTACAGCTCCGAGAGAGTATGAATTTGACCTGAGGAAAGTTCTGCTTC 1140  
1141 TCAGTCAGGTTTAAAGCGTACCCAGGAATCCGATGCACTGGATCTTCTCTCAAGCCTCA 1200  
1141 TCAGTCAGGTTTAAAGCGTACCCAGGAATCCGATGCACTGGATCTTCTCTCAAGCCTCA 1200  
1201 TTTTCTTGTGAACAGAGAGGCTGGAGGTGGGTACAGCATATCTAAATTTTTCGATCAT 1260

1201 TTTTCTTGTGAACAGAGAGGCTGGAGATGGGTACAGCATATCTAAATTTTTCGATCAT 1260  
1261 AAGAACAGCCAGGAGAGTACATATTTCTATGCGAATAATGATGACGCCAGTTCACCAAA 1320  
1261 AAGAACAGCCAGGAGAGTACATATTTCTATGCGAATAATGATGACGCCAGTTCACCAAA 1320  
1321 ATGTTTCACTGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAATGCTTCAGCCAGCCAG 1380  
1321 ATGTTTCACTGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAATGCTTCAGCCAGCCAG 1380  
1381 GCGTCTGTTTCTCTGATGCTACCCGCTACCTTCTTGGACCTCGAAGAAAGTGTTCGAC 1440  
1381 GCGTCTGTTTCTCTGATGCTACCCGCTACCTTCTTGGACCTCGAAGAAAGTGTTCGAC 1440  
1441 AAATCTCCAAATTCACCGAGGAAATCCCAAGAGAGTGTGGAATAAAAGGCTTAACAGA 1500  
1441 AAATCTCCAAATTCACCGAGGAAATCCCAAGAGAGTGTGGAATAAAAGGCTTAACAGA 1500  
1501 AAAGTGTTCGCGCAGTGGGTGTGAGCAGTACTCTAAATATGAGTGAGCCCGGAAAGGG 1560  
1501 AAAGTGTTCGCGCAGTGGGTGTGAGCAGTACTCTAAATATGAGTGAGCCCGGAAAGGG 1560  
1561 CTTCTGCTCAAAATGCTGCTGCTACAAATTTCTATGGGCAGCTCTTGGAAACCATCTTTT 1620  
1561 CTTCTGCTCAAAATGCTGCTGCTACAAATTTCTATGGGCAGCTCTTGGAAACCATCTTTT 1620  
1621 AACTCACAGGCCCTTCCCTTTTCATCCAGACAACTCTCTTCTATGCGACCATTTGG 1680  
1621 AACTCACAGGCCCTTCCCTTTTCATCCAGACAACTCTCTTCTATGCGACCATTTGG 1680  
1681 CTCTGCTCTCCCTTCTCATTTGTTTCTCATTTGTTGATCTGCCCAAAATACAAAGGAA 1740  
1681 CTCTGCTCTCCCTTCTCATTTGTTTCTCATTTGTTGATCTGCCCAAAATACAAAGGAA 1740  
1741 TTTAGGTACAGAGTCACTGCTGAGTATCCAGGTGACTGGCCCCCTGGATTAACAGTAC 1800  
1741 TTTAGGTACAGAGTCACTGCTGAGTATCCAGGTGACTGGCCCCCTGGATTAACAGTAC 1800  
1801 TTTACGTTGACTTCAGGACCTATGAATATGAACCTTAAGTGGAGTTCGCGAGAGAAC 1860  
1801 TTTACGTTGACTTCAGGACCTATGAATATGAACCTTAAGTGGAGTTCGCGAGAGAAC 1860  
1861 TTAGAGTTTGGAGAGTCTGGGCTCTGGGCTTTCGGAGGCTGATGAACCCACGCGC 1920  
1861 TTAGAGTTTGGAGAGTCTGGGCTCTGGGCTTTCGGAGGCTGATGAACCCACGCGC 1920  
1921 TATGGCATTTAGTAAACCGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAGAGAA 1980  
1921 TATGGCATTTAGTAAACCGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAGAGAA 1980  
1981 GCTGACGCTGTGAAAAAGAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA 2040  
1981 GCTGACGCTGTGAAAAAGAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA 2040  
2041 CACATGACAACTGCTGAATCTGCTGGGGGATGCACTGTCAGGCGCAGTGTACTTTG 2100  
2041 CACATGACAACTGCTGAATCTGCTGGGGGATGCACTGTCAGGCGCAGTGTACTTTG 2100  
2101 ATTTTGAATATTTGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTG 2160  
2101 ATTTTGAATATTTGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTG 2160  
2161 TTTTCAAGGACATGGACAGAGATTTTAAAGGAACTAAATTTTCACTTCTTCTTCTTCTTCT 2220  
2161 TTTTCAAGGACATGGACAGAGATTTTAAAGGAACTAAATTTTCACTTCTTCTTCTTCTTCT 2220  
2221 CAGGCACATTCAAATTCAGGAGTCTGTTTCAAGAGGTTTCACTTACCCGCGCTTG 2280  
2221 CAGGCACATTCAAATTCAGGAGTCTGTTTCAAGAGGTTTCACTTACCCGCGCTTG 2280  
2281 GATCAGCTCTCAGGTTTCAATGGGATTTCAATTTCAATTTCAATTTCAATTTCAATTTCAAT 2340

Db 2281 GATCAGCTCTCAGGGTTCAATGGGAATTCATTTCTCAAGATGAGATTGAATGAA 2340  
QY 2341 AACCAAGAGGCTGGCAGAGAGAGAGAGAGAGATTTGAACGCTGCTGAGCTTTGAAGAC 2400  
Db 2341 AACCAAGAGGCTGGCAGAGAGAGAGAGAGATTTGAACGCTGCTGAGCTTTGAAGAC 2400  
QY 2401 CTCCTTTGCTTTGCTGACCAAGTGGCCAAAGGATGGAATTCCTGGAAGTTCAAGTCTGT 2460  
Db 2401 CTCCTTTGCTTTGCTGACCAAGTGGCCAAAGGATGGAATTCCTGGAAGTTCAAGTCTGT 2460  
QY 2461 GTCCACAGAGACTGCGCAGCAGGATGTTGTTGTCACCCACGGGAAGGTGTTGAAGATC 2520  
Db 2461 GTCCACAGAGACTGCGCAGCAGGATGTTGTTGTCACCCACGGGAAGGTGTTGAAGATC 2520  
QY 2521 TGTGACTTTGGACTGCGCCGAGACATCTCGAGGACTCCAGCTACCTGCTCAGGGGCAAC 2580  
Db 2521 TGTGACTTTGGACTGCGCCGAGACATCTCGAGGACTCCAGCTACCTGCTCAGGGGCAAC 2580  
QY 2581 GCACGGCTGCCGTTGAAGTGGATGGCACCAGAGCTTATTTGAAGGGATCTACACATC 2640  
Db 2581 GCACGGCTGCCGTTGAAGTGGATGGCACCAGAGCTTATTTGAAGGGATCTACACATC 2640  
QY 2641 AAGAGTGAGCTGCTGCTTACGGGATCTCTCTCGGAGATATTTTCACTGGGTGTGAAC 2700  
Db 2641 AAGAGTGAGCTGCTGCTTACGGGATCTCTCTCGGAGATATTTTCACTGGGTGTGAAC 2700  
QY 2701 CTTTACCTGCGCATTCCTGCTGACGCTAACTTCTTATAAATGATTCAGAGTGGATTTAA 2760  
Db 2701 CTTTACCTGCGCATTCCTGCTGACGCTAACTTCTTATAAATGATTCAGAGTGGATTTAA 2760  
QY 2761 ATGAGAGCATTCTATGCCACAGAGGATATCTTTGTAATGCAATCCCTGCTGGGCT 2820  
Db 2761 ATGAGAGCATTCTATGCCACAGAGGATATCTTTGTAATGCAATCCCTGCTGGGCT 2820  
QY 2821 TTTGACTCAAGAGCGGCTTCCCTTCCCAACCTGACTTCAATTTTAGTAGTGTGAGCTG 2880  
Db 2821 TTTGACTCAAGAGCGGCTTCCCTTCCCAACCTGACTTCAATTTTAGTAGTGTGAGCTG 2880  
QY 2881 GCAGAGCAGAGAGCATGTATCAGACATCCATCCATCCATCCATCCATCCATCCATCC 2940  
Db 2881 GCAGAGCAGAGAGCATGTATCAGACATCCATCCATCCATCCATCCATCCATCCATCC 2940  
QY 2941 CAGCAGAGAGCGGCTCAGAGCCGCTGCGCCACAGCCGCTGAGTGAAGATTCACAGAA 3000  
Db 2941 CAGCAGAGAGCGGCTCAGAGCCGCTGCGCCACAGCCGCTGAGTGAAGATTCACAGAA 3000  
QY 3001 AGAAGTTAGCAGAGGCTTTGGACCCCGCCACCTAGCAGCTGTAGACCCGAGGCA 3060  
Db 3001 AGAAGTTAGCAGAGGCTTTGGACCCCGCCACCTAGCAGCTGTAGACCCGAGGCA 3060  
QY 3061 AGATTAGCCTGCTGAGAGAGGCGCTCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 3120  
Db 3061 AGATTAGCCTGCTGAGAGAGGCGCTCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 3120  
QY 3121 AGATGCTGTCTGCCATTCTCAAGTGAATTTCTATAAATCAAACTCTCTCGCAG 3180  
Db 3121 AGATGCTGTCTGCCATTCTCAAGTGAATTTCTATAAATCAAACTCTCTCGCAG 3180  
QY 3181 GCGGGAGAGCAATATGAGACTTTGTTGGTGGAGCCGCTACCTGCGGGGCTTTTCCACG 3240  
Db 3181 GCGGGAGAGCAATATGAGACTTTGTTGGTGGAGCCGCTACCTGCGGGGCTTTTCCACG 3240  
QY 3241 AGCTTGAGGGAGAGCCTATCTGAAATATAGTATATTTCTGTAATACGTGAACAA 3300  
Db 3241 AGCTTGAGGGAGAGCCTATCTGAAATATAGTATATTTCTGTAATACGTGAACAA 3300  
QY 3301 ACCAAACCCCTTTTCTGAAGAGAGCTAAATATGATTTTAAATCTATGTTTAA 3360  
Db 3301 CCAAAACCCCTTTTCTGAAGAGAGCTAAATATGATTTTAAATCTATGTTTAA 3360  
QY 3361 AATACTATGATCTTTTCTATCTATTTAGTGAATATTTTATGATGGAATTAACCTTC 3420  
Db 3361 AATACTATGATCTTTTCTATCTATTTAGTGAATATTTTATGATGGAATTAACCTTC 3420

QY 3421 TACTGTAAAAA 3453  
Db 3421 TACTGTAAAAA 3453

## RESULT 11

AAD42483  
ID AAD42483 standard; cDNA; 3440 BP.

XX AAD42483;

XX 15-NOV-2002 (first entry)

DE Murine receptor protein tyrosine kinase, flk-2 cDNA.

XX Murine; receptor protein tyrosine kinase; pTK; haematopoietic cell;  
KW growth factor; therapeutic; macrocytic anaemia; aplastic anaemia;  
KW bone marrow damage; cancer; chemotherapy; radiation; osteopathic;  
XX flk-2; gene; ss.

OS Mus sp.

XX Key Location/Qualifiers  
EH CDS 31..3009

FT /tag= a  
FT /product= "Murine flk-2 protein"

FT sig\_peptide 31..111

FT mat\_peptide /tag= b  
112..3006

FT /product= "Murine mature flk-2 protein"

FT misc\_feature 112..1662

FT /tag= d

FT /note= "Encodes extracellular domain"

FT /tag= e  
1663..1722

FT /note= "Encodes transmembrane domain"

FT misc\_feature 1723..3006

FT /tag= f

FT /note= "Encodes intracellular domain"

XX US2002072077-A1.

PN 13-JUN-2002.

XX 31-JUL-2001; 2001US-0919408.

XX 19-NOV-1992; 92US-0977451.

XX 30-APR-1993; 93US-0055269.

XX 31-OCT-1994; 94US-0252498.

XX 15-FEB-1996; 96US-0601891.

XX 10-FEB-1998; 98US-0021324.

XX 10-DEC-1998; 98US-0208786.

XX 02-APR-1991; 91US-0679666.

XX 28-JUN-1991; 91US-0728913.

XX 15-NOV-1991; 91US-0793065.

XX 24-DEC-1991; 91US-0813593.

XX (LEMI/) LEMISCHKA I R.

XX Lemischka IR;

XX WPI; 2002-607237/65.

XX P-PSDB; AAE25818.

New protein tyrosine kinase expressed in primitive hematopoietic cells (HC) and not expressed in mature HC, and ligands for the protein, for stimulating proliferation of primitive hematopoietic stem cells.

Claim 6; Page 15-19; 64pp; English.

The present invention relates to receptor protein tyrosine kinase (pTK)

expressed in primitive haematopoietic cells and not expressed in mature haematopoietic cells, polynucleotides encoding such proteins and ligands for the protein. Ligands which bind to pRk of the invention are useful for stimulating the proliferation and/or differentiation of primitive mammalian haematopoietic stem cells. The receptor pRk molecules are useful for stimulating the self renewal of the totipotent haematopoietic stem cell and to stimulate the development of all cells of haematopoietic system both in vitro and in vivo. The ligands for the receptors act as haematopoietic growth factors. The ability of the ligands to stimulate proliferation of stem cells both in vitro and in vivo has important therapeutic applications such as treating humans whose primitive stem cells do not sufficiently undergo self-renewal. It is also useful in conditions that occur when defects in haematopoietic stem cells or their related growth factors depress the number of white blood cells such as macrocytic and aplastic anaemia or bone marrow damage resulting from cancer chemotherapy and radiation. The present sequence is a cDNA encoding murine receptor pRk, flk-2.

XX  
SQ Sequence 3440 BP; 943 A; 819 C; 874 G; 804 T; 0 other;

Query Match 97.2%; Score 3357; DB 24; Length 3440;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 3440: Conservative 0; Mismatches 0; Indels 13;

[illegible]

Db	721	AGATGCTGTGCTAGAAATGCACTGGGCGCGAAATGCACCAAGCTGTTACACCATAGATCTA	780
Qy	781	AAACAGGCTCCTCAGAGCACACTGCCCCAGTTATTCTCTGAAAGTGGGGGAAACCCCTTGTTGG	840
Db	781	AAACAGGCTCCTCAGAGCACACTGCCCCAGTTATTCTCTGAAAGTGGGGGAAACCCCTTGTTGG	840
Qy	841	ATCAGGTGTAAAGCCATPCCATGTGAACCATATGGATTCGGGCTCACTCGGAGCTGGAAGAC	900
Db	841	ATCAGGTGTAAAGCCATPCCATGTGAACCATATGGATTCGGGCTCACTCGGAGCTGGAAGAC	900
Qy	901	AAAGCCCTGAGAGGGCGAGCTACTTTGAGATGAGTACCTACTTCCACAAACAGGACCATTG	960
Db	901	AAAGCCCTGAGAGGGCGAGCTACTTTGAGATGAGTACCTACTTCCACAAACAGGACCATTG	960
Qy	961	ATTCCGATTTCTTTGGCTTTTGTGTCTTCGCTGGGAAGAACGACACCGGATATTACACC	1020
Db	961	ATTCCGATTTCTTTGGCTTTTGTGTCTTCGCTGGGAAGAACGACACCGGATATTACACC	1020
Qy	1021	TGCTCTTTCTCAAAGCACCCAGCCAGCTCAGCGTTTGGTGACCATCTCTAGAAAAAGGGTTT	1080
Db	1021	TGCTCTTTCTCAAAGCACCCAGCCAGCTCAGCGTTTGGTGACCATCTCTAGAAAAAGGGTTT	1080
Qy	1081	ATAAACGCTACCAAGTCCGAAGAAGATATGAATTTGACCCGTAAGAAAGTTCTGCTTC	1140
Db	1081	ATAAACGCTACCAAGTCCGAAGAAGATATGAATTTGACCCGTAAGAAAGTTCTGCTTC	1140
Qy	1141	TCAGTCAGGTTTAAAGCGTACCCACGAATCCGATGCACGTGGATCTCTCTCAAGCCCTCA	1200
Db	1141	TCAGTCAGGTTTAAAGCGTACCCACGAATCCGATGCACGTGGATCTCTCTCAAGCCCTCA	1200
Qy	1201	TTTTCTTTGTGAAACAGAGAGGCTGGAGATGGGTACAGCATATCTAAATTTTGGCATCAT	1260
Db	1201	TTTTCTTTGTGAAACAGAGAGGCTGGAGATGGGTACAGCATATCTAAATTTTGGCATCAT	1260
Qy	1261	AAGAAACAGCCAGGAGATACATTTCTTATGCAGAAAAATGATGAGCCAGCTTCAACCAA	1320
Db	1261	AAGAAACAGCCAGGAGATACATTTCTATGCAGAAAAATGATGAGCCAGCTTCAACCAA	1320
Qy	1321	ATGTTTCAGCTGTAATATAAGAAAGAAACCTCAAGTGTCTAGCAATGCTCTCAGCAGCCAG	1380
Db	1321	ATGTTTCAGCTGTAATATAAGAAAGAAACCTCAAGTGTCTAGCAATGCTCTCAGCAGCCAG	1380
Qy	1381	CGCTCTGTTCTCTGATGGCTACCCGCTACCCCTCTTTGGACCTCGAAAGAGTGTTCGGAC	1440
Db	1381	CGCTCTGTTCTCTGATGGCTACCCGCTACCCCTCTTTGGACCTCGAAAGAGTGTTCGGAC	1440
Qy	1441	AAATCTCCCAATTGCACGGAGGAAATCCCAAGAGGAGTTTGGAAATAAAAGGCTTAAACAGA	1500
Db	1441	AAATCTCCCAATTGCACGGAGGAAATCCCAAGAGGAGTTTGGAAATAAAAGGCTTAAACAGA	1500
Qy	1501	AAAGTGTGTGGCCAGTGGGTGTCGAGCAGTACTCTAAATATGAGTGAAGCCGGGAAAGGG	1560
Db	1501	AAAGTGTGTGGCCAGTGGGTGTCGAGCAGTACTCTAAATATGAGTGAAGCCGGGAAAGGG	1560
Qy	1561	CTTCTGGTCAAAATGCTGTGCGTACAAATCTATGGGCACGCTCTTCGGAACCATCTTTTAA	1620
Db	1561	CTTCTGGTCAAAATGCTGTGCGTACAAATCTATGGGCACGCTCTTCGGAACCATCTTTTAA	1620
Qy	1621	AACTCACCAGGCCCTTCCCTTTTCATCCAAAGACAATCTCTCTTCTATGCGACCATTTGGG	1680
Db	1621	AACTCACCAGGCCCTTCCCTTTTCATCCAAAGACAATCTCTCTTCTATGCGACCATTTGGG	1680
Qy	1681	CTCTGTCTCCCTTCATGTTGTGTTCTCATTTGTGTTGATCTGCCACAAATACAAAAAGCAA	1740
Db	1681	CTCTGTCTCCCTTCATTTGTTGTTCTCATTTGTGTTGATCTGCCACAAATACAAAAAGCAA	1740
Qy	1741	TTTAGGTACGAGAGTCAAGTGCAGATCATCCAGTGAAGTGCAGTGGGATTAACGAGTAC	1800
Db	1741	TTTAGGTACGAGAGTCAAGTGCAGATCATCCAGTGAAGTGCAGTGGGATTAACGAGTAC	1800
Qy	1801	TTCTACGTTTGACCTTCAGGAGTACTGAATATGACCTTAAGTGGGAGTTCCCGAGAGAGAAC	1860





QY 1921 TATGGCAATTAGTAAACCGGAGTCTCAATTACAGTGGCGGTGAGAGTCTAAAGAGAA 1980  
 Db 1972 TATGGCAATTAGTAAACCGGAGTCTCAATTACAGTGGCGGTGAGAGTCTAAAGAGAA 2031  
 QY 1981 GCTGACAGCTGTGAAAGAGCTCTCATGTGCGAGCTCAAAATGATGACCCACTGGGA 2040  
 Db 2032 GCTGACAGCTGTGAAAGAGCTCTCATGTGCGAGCTCAAAATGATGACCCACTGGGA 2091  
 QY 2041 CACCATGACACATCGTAATCTGCTGGGGGATGACACTGTGAGGCGGAGTGTACTTG 2100  
 Db 2092 CACCATGACACATCGTAATCTGCTGGGGGATGACACTGTGAGGCGGAGTGTACTTG 2151  
 QY 2101 ATTTTGAATATTGTTGCTATGTGACCTCTCACTACCTACCTAAGATTAAGAGAGAG 2160  
 Db 2152 ATTTTGAATATTGTTGCTATGTGACCTCTCACTACCTAAGATTAAGAGAGAG 2211  
 QY 2161 TTTCACAGACATGACAGAGATTTTAAAGAAACATAATTTCACTTACCTTACTTTC 2220  
 Db 2212 TTTCACAGACATGACAGAGATTTTAAAGAAACATAATTTCACTTACTTTC 2271  
 QY 2221 CAGGCACATCAAAATTCAGACATGCTGTTTCAAGAGATTCAGTTACACCGCCCTTG 2280  
 Db 2272 CAGGCACATCAAAATTCAGACATGCTGTTTCAAGAGATTCAGTTACACCGCCCTTG 2331  
 QY 2281 GATCAGCTCTCAGGTTTCAATGGGAATTCATTTTCAAGATTCAGTTACAGATTCGA 2340  
 Db 2332 GATCAGCTCTCAGGTTTCAATGGGAATTCATTTTCAAGATTCAGTTACAGATTCGA 2391  
 QY 2341 AACCAAGAGAGCTGACAGAGAGAGAGAGATTTGAACGCTGCTGAGTTTGAAGAC 2400  
 Db 2392 AACCAAGAGAGCTGACAGAGAGAGAGAGATTTGAACGCTGCTGAGTTTGAAGAC 2451  
 QY 2401 CTCCTTTGCTTTCGTTCCAGTGGCCAAAGGATTCCTGAGTTCAGTCTGT 2460  
 Db 2452 CTCCTTTGCTTTCGTTCCAGTGGCCAAAGGATTCCTGAGTTCAGTCTGT 2511  
 QY 2461 GTCCACAGAGACTGCGACGAGGATGTTGTTCAACCCAGGAGGTTGTTGAAGATC 2520  
 Db 2512 GTCCACAGAGACTGCGACGAGGATGTTGTTCAACCCAGGAGGTTGTTGAAGATC 2571  
 QY 2521 TGTGACTTTGACTGCGCCGAGACATCTGAGGACTCCAGTACGCTGCTGAGGGCAAC 2580  
 Db 2572 TGTGACTTTGACTGCGCCGAGACATCTGAGGACTCCAGTACGCTGCTGAGGGCAAC 2631  
 QY 2581 GCA CGGCTGCGGTTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGATCTACAAATC 2640  
 Db 2632 GCA CGGCTGCGGTTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGATCTACAAATC 2691  
 QY 2641 AAGAGTGACGCTGCTGCTACGGCATCTTCTCTGGGAGATTTTCACTGGGTGTGAAC 2700  
 Db 2692 AAGAGTGACGCTGCTGCTACGGCATCTTCTCTGGGAGATTTTCACTGGGTGTGAAC 2751  
 QY 2701 CCTTACCTGCGATTCCTGCTGACGCTAACTTCTATAAATGATTCAGAGTGGATTTAA 2760  
 Db 2752 CCTTACCTGCGATTCCTGCTGACGCTAACTTCTATAAATGATTCAGAGTGGATTTAA 2811  
 QY 2761 ATGAGAGAGCATTCTATGCCACAGAGGATATCTTTGTAATGCAATCTCTGCTGGGT 2820  
 Db 2812 ATGAGAGAGCATTCTATGCCACAGAGGATATCTTTGTAATGCAATCTCTGCTGGGT 2871  
 QY 2821 TTTGACTCAAGGAGCGGCGCATCTTCTCCCACTGACTTCAATTTTGAAGTGTGAGTGT 2880  
 Db 2872 TTTGACTCAAGGAGCGGCGCATCTTCTCCCACTGACTTCAATTTTGAAGTGTGAGTGT 2931  
 QY 2881 GCAGAGCAGAGAGAGC-----ATGATATCAGAACATCCATCC 2917  
 Db 2932 GCAGAGCAGAGAGAGC-----ATGATATCAGAACATCCATCC 2991  
 QY 2918 ATCTACCAAAACAGCGGCGCCTCAGCAGAGAGCGGCTCAGAGCC-CAGTGGCCACAG 2976  
 Db 2992 ATCTACCAAAACAGCGGCGCCTCAGCAGAGAGCGGCTCAGAGCCCGCCATCGCCACAG 3051

QY 2977 GGCAGGTGAGATTTCACAGAGAAAGATTAGCGAGAGGCGCTTGGACCCCGCCACCT 3036  
 Db 3052 GGCAGGTGAGATTTCACAGAGAAAGATTAGCGAGAGGCGCTTGGACCCCGCCACCT 3111  
 QY 3037 AGCAGGCTGTAGACCGCAGAGCCAAAGATTAGCTCGCTCTGAGAGAGCGCCCTACAGG 3096  
 Db 3112 AGCAGGCTGTAGACCGCAGAGCCAAAGATTAGCTCGCTCTGAGAGAGCGCCCTACAGG 3171  
 QY 3097 CGTTGCTCGCTGAGACTTTTCTTAGATGCTGCTGCAATTAATCTCAAAAGTGAATCTTAT 3156  
 Db 3172 CGTTGCTCGCTGAGACTTTTCTTAGATGCTGCTGCAATTAATCTCAAAAGTGAATCTTAT 3231  
 QY 3157 AAAATCAAACCTCTCTGCAAGCGGAGAGCCAAATAATAGACTTTGTTGGTGAGCCC 3216  
 Db 3232 AAAATCAAACCTCTCTGCAAGCGGAGAGCCAAATAATAGACTTTGTTGGTGAGCCC 3291  
 QY 3217 GCCTACCTGGGGGCGCTT-----CCACGAGCTTGGGGGAAAGCCATGTATCTGAA 3269  
 Db 3292 GCCTACCTGGGGGCGCTT-----CCACGAGCTTGGGGGAAAGCCATGTATCTGAA 3351  
 QY 3269 ATATAGTATATTTCTTGAATACTGTAACAAACCCGCTTTTGTCTAAGGAAAG 3328  
 Db 3352 ATATAGTATATTTCTTGAATACTGTAACAAACCCGCTTTTGTCTAAGGAAAG 3411  
 QY 3329 CTAAATATGATTTTAAATCTATGTTTAAATACTATGTAACCTTTTCACTATTTA 3388  
 Db 3412 CTAAATATGATTTTAAATCTATGTTTAAATACTATGTAACCTTTTCACTATTTA 3471  
 QY 3389 GTGATATATTTATGATGGAATAAATCTTCTACTGTAATAAAAAA 3438  
 Db 3472 GTGATATATTTATGATGGAATAAATCTTCTACTGTAATAAAAAA 3521

RESULT 13

AAQ81013  
 ID AAQ81013 standard; cDNA; 3501 BP.  
 XX AAQ81013;  
 AC AAQ81013;  
 XX 25-MAR-2003 (updated)  
 DT 18-AUG-1995 (first entry)  
 DE Flk2 receptor protein-tyrosine-kinase cDNA.  
 XX  
 KW Human Flk2; receptor protein-tyrosine-kinase; primitive  
 KW hematopoietic cell; fetal liver kinase; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 CDS 58..3039  
 FT /\*tag= a  
 FT /product= Flk2 receptor protein-tyrosine-kinase  
 FT 58..138  
 FT /\*tag= b  
 FT 139..3036  
 FT mat\_peptide /\*tag= c  
 XX  
 PN WO9500554-A2.  
 XX  
 PD 05-JAN-1995.  
 XX  
 PF 17-JUN-1994; 94WO-US06944.  
 XX  
 PR 18-JUN-1993; 93US-0080244.  
 PR 21-JUN-1993; 93US-0081508.  
 PR 23-NOV-1993; 93US-0157490.  
 XX  
 PA (UYPR-) UNIV PRINCETON.  
 XX  
 PI Lemischka IR;  
 XX  
 WP1; 1995-052014/07.





1868 TTGGGAAGGTCTCTGGGCTCTGGGCTTTTCGGAGGTGATGAACCCAGCGCTTATGGCA 1927  
 1892 TTGGGAAGGTACTAGCATCAGGTGCTTTTGGAAAGGTGATGAACCCAGCGCTTATGGAA 1951  
 1928 TTAGTAAACGGGAGTCTCAATTCAGTGGCGGTGAAGTCTTAAAGAGAAAGCTGACA 1987  
 1952 TTAGCAAAACAGGAGTCTCAATCCAGGTTCGCGTCAAAATGCTGAAAGAAAGACGACA 2011  
 1988 GCTGTGAAAGAAAGCTCTCATCTCGGAGCTCAAAATGATGATCCACCTGGGACACATG 2047  
 2012 GCTCTGAAGAGAGGCACTCATGTGCAAACTCAAGATGATGATCCAGCTGGGAAGCCAG 2071  
 2048 ACAACATCTGAATCTGCTGGGGCAATGCACTGTCAAGGCGAGTGTACTTGAATTTTG 2107  
 2072 AGAATATTGTGAACCTGCTGGGGGTGCACTGTCAAGACCAATTTACTTTGATTTTG 2131  
 2108 ATATATTGTGCTATGGTGAACCTCTCACTAAGAAAGTAAAGAGAGAAAGTTTCA 2167  
 2132 AATACTGTGTGCTATGGTGAATCTCTCACTAAGAAAGTAAAGAGAGAAATTTCA 2191  
 2168 GGACATGACAGAGATTTTAAAGGAACATAATTTCACTTCTTACCTACTTTCCAGGCAC 2227  
 2192 GGACTTGGACAGAGATTTTCAAGGAACATAATTTCACTTCTTACCTTCTTCCAACTAC 2251  
 2228 ATTCAAAATTCAGCATGCTGCTTCAAGAGAGTTCAGTTACACCGCGCTTGGATCAGC 2287  
 2252 ATCCAAATTCAGCATGCTGCTTCAAGAGAGTTCAGTTACACCGCGCTTGGATCAGC 2311  
 2288 TCTCAGGTTCAATGGGAATTCATTTCAAGATGAGATTCAGTTCAATGAAACCCAGA 2347  
 2312 TCTCAGGCTTCATGGGAATTCATTTCACTTCAAGATGAGATTCAGTTCAATGAAACCCAA 2371  
 2348 AGAGCTTGGCAGACAGAGAGAGAGATTTGAAGCTGCTGAGCTTCAAGACCTCTTT 2407  
 2372 AAGGCTG-----GAAGAGAGAGAGACTTGAATGTGCTTACATTTCAAGATCTCTTT 2425  
 2408 GCTTTGGCTACCAAGTGCCAAAGCA TGGAAATTCCTGGAGTCAAGTCTCAAGTCTGCTCCACA 2467  
 2426 GCTTTGCATATCAAGTTTGCCTCAAGAGATGGAATTTCTGGAATTTAAGTCTGCTTCA 2485  
 2468 GAGACCTGGCAGCAGGAATGTGTTGTCACCCAGGAAGTGTGAAGATCTGTGACT 2527  
 2486 GAGACCTGGCGCCAGCAAGCTGCTGTCACCCAGGAAGTGTGAAGATCTGTGACT 2545  
 2528 TTGAGCTTGGCGCCAGACATCTGAGCCACTCCAGCTAGCTGCTCAGGGGCAACGACGGC 2587  
 2546 TTGATTTGGCTCAGATCATGATGATGATTCCTCAACTATGTTGTCAGGGGCAATGCCGTC 2605  
 2588 TGCCGCTGAAGTGGATGGCAGCCGAGAGCTTATTTGAAGGATCTACACATCAAGAGTG 2647  
 2606 TGCCCTGTAATATGATGGCCCGCGAAGCTGTTTGAAGGATCTACACATTAAGAGTG 2665  
 2648 ACCTGTGCTTACGGCATCTTCTCTGGGAGATATTTTCACTGGGTGTCAACCTTTACC 2707  
 2666 ATGCTCTGGTCATATGAATATTAATCTGTGGGAATCTTCTCACTTGGTGTGAATCTTACC 2725  
 2708 CTGGCATCTCTGCGAGCTAACTTCTAATACTGATTCAGTGTGAATTTAAATAGGAGC 2767  
 2726 CTGGCAATCCGTTGATGTAACTTCTCAAACTGATTCAAATGATTTAAATGGATC 2785  
 2768 AGCCATTTCTATGCCAGAGGATATCTTCTGATGATGATGATGATGATGATGATGATGATG 2827  
 2786 AGCCATTTTATGCTACAGAGAAATATACATTAATGATGATGATGATGATGATGATGATG 2845  
 2828 CAAGGAAGCGGCATCTTCCCAACCTGACTTCAATTTTAGGATGCTGATGCTGCGAGAGG 2887  
 2846 CAAGGAAGCGGCATCTTCCCTAAATTGACTTCTGTTTGGATGCTGATGCTGCGAGAGG 2905  
 2888 CAGAGAAG-----CATGATCAGAAACATCCATCCATCTACC 2924  
 2906 CAGAGAAGCGATGATCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 2965

2925 AAAACAGGCGGCGCTTACAGCAGAGAGAGCGGCGCTCAGAGC-CCAGTCCGCCAGAGCGCCAGG 2983  
 2966 AAAACAGGCGGCGCTTACAGCAGAGAGAGCGGCGCTTACAGCAGAGAGAGCGGCGCTCAGG 3025  
 2984 T-GAAGATTACAG 3038  
 3026 TCGAGATTCTGAG 3085  
 3039 CAGGCTGTAGACCGCAG 3096  
 3086 CAGGCTGTAGATTACCAAAACAAGATTAATTTTCACTAAAGAGAGAGAGAGAGAGAGAGAG 3145  
 3097 CGTTGCTGCTGGAGCTTTTCTCTAGATGCTGTCTGCCATTACTC-----CAAAAGTGA 3149  
 3146 TCGTCTGCTTCCAGAGAGCTTTTCTCTAGAGCGCTGCGTGTACTCTTTGTTTCAAGGGA 3205  
 3150 CTTCTATATAAAATCAAACTCTCTCTCCACAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 3209  
 3206 CTTTGTAAATCAAAATCATCTGTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3265  
 3210 TGAGCCCGCTACCTCGCTGGGGGCTTTCCAGAGCTTTCAGGGGAGAGAGAGAGAGAGAGAG 3269  
 3266 GAGCATTGATCTGCAATCAAGGCTTCTCAGCGCGCTTGTAGTGAATTTGTACCTGAAG 3325  
 3270 TATAGTATATTCTTGTAAATACGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3329  
 3326 TACAGTATATTCTGTAAATACATAAACA-----AGCATTTTGTAGGAGAGAGAGAGAG 3378  
 3330 TAAATATGATTTTAAATACTTGTGTTTAAATACTATGTAATTTTCACTTATTTAG 3389  
 3379 TAAATATGATTTT-----AAGTCTATGTTTAAATAATATGTAATTTTCACTTATTTAG 3436  
 3390 TGATATATTTTATGATGGAATAAATCTTCTCTCTGTAATAAATAAATAAATAAATAAATAA 3449  
 3437 TGATATATTTTATGAGTGGAGTAAATTTTCTACAGAGAGAGAGAGAGAGAGAGAGAGAG 3496  
 3450 AAAA 3453  
 3497 AAAA 3500

RESULT 14  
 AAQ79069  
 ID AAQ79069 standard; cDNA; 3501 BP.  
 AC AAQ79069;  
 XX  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 04-JUL-1995 (first entry)  
 DE Human flk-2 cDNA.  
 XX  
 KW Fetal liver kinase-2; flk-2; protein tyrosine-kinase receptor;  
 KW hematopoiesis; stem cell; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key  
 CDS 58..3039  
 FT sig\_peptide /\*tag= a  
 FT mat\_peptide /\*tag= b  
 FT 139..3036 /\*tag= c  
 XX  
 PN US5367057-A.  
 XX  
 PD 22-NOV-1994.  
 XX  
 PF 30-APR-1993; 93US-0055269.  
 XX  
 PR 02-APR-1991; 91US-0679666.  
 PR 28-JUN-1991; 91US-0728913.

PR 15-NOV-1991; 91US-0793065.  
 PR 24-DEC-1991; 91US-0813593.  
 PR 26-JUN-1992; 92US-0906397.  
 PR 12-NOV-1992; 92US-0975049.  
 PR 19-NOV-1992; 92US-0977451.  
 PR 30-APR-1993; 93US-0055269.  
 XX (UYPR-) UNIV PRINCETON.  
 PA Lemischka IR;  
 XX WPI; 1995-005894/01.  
 DR P-PSDB; AAR67536.  
 XX Murine flk-2 receptor protein tyrosine kinase - used to stimulate proliferation and/or stimulation of primitive mammalian haematopoietic stem cells in vitro or in vivo.  
 PT Disclosure; Fig. 2A-1F; 69pp; English.  
 XX cDNAs encoding receptor protein tyrosine-kinases, mouse fetal liver kinase-2 (flk-2), human flk-2 and mouse flk-1 are given in AAQ79068-70, CC respectively, and the deduced amino acid sequences in AAR67535-37, CC respectively.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX SQ Sequence 3501 BP; 1068 A; 709 C; 784 G; 940 T; 0 other;  
 Query Match 64.5%; Score 2227.2; DB 16; Length 3501;  
 Best Local Similarity 79.6%; Pred. No. 0;  
 Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;  
 QY 8 GGCTACCGCGCTCCGAGGCCATCGCGGCGTTGGCGCAGCGCAGCGCGCGGCTGC 67  
 DB 35 GGGGACCCCGGCTCCGAGGCCATCGCGGCGTTGGCGGCGCGCGCGCGCGGCTGC 91  
 QY 68 TGCTGCTGTTGTTGTCAGTAATGATCTTGGACCGTTTACCAACCAAGACCTGCGCTG 127  
 DB 92 CGCTGCTGCTGTTGTTGTCAGTAATGATCTTGGGACTATTACAAATCAAGATCTGCGCTG 151  
 QY 128 TGATCAAGTGTTTAAATCAGTCATGAGAACATGGCTCATCAGCGGGAAGCCATCAT 187  
 DB 152 TGATCAAGTGTTTAAATCAGTCATGAGAACATGGCTCATCAGTCGGAAGTATCAT 211  
 QY 188 CGTACCGAATGTGGGAGGATCCCGAGAGACCTCCAGTGATCCCGAGCGCGCCAGAGTG 247  
 DB 212 CATATCCCATGGTATCAGATCCCGGAGACCTCGGGTGGGTGAGACCCAGAGCT 271  
 QY 248 AAGGACCGGTATGAGAGCGGACCGTGAGGTGGCGAGTCTGGGTCCATCACCCTGC 307  
 DB 272 CAGGACAGTGTACGAAGCTGCGCTGTGGAAGTGGATGTATCTGCTTCCATCACACTGC 331  
 QY 308 AAGTGCAGTCCGACCCAGGGACCTTCTGCTCTGGGTCTTAAAGCACAGCTCCC 367  
 DB 332 AAGTGCAGTCCGACCCAGGGACATTTCTGCTCTGGGTCTTAAAGCACAGCTCCC 391  
 QY 368 TGGGCTGCGACCGCACTTTGATTTACAAAACAGAGGAATCGTTTCCATGGCCATCTTGA 427  
 DB 392 TGAATTGCGACCCATTTGATTTACAAAACAGAGGAGTGTTCCTATGTCATTTGA 451  
 QY 428 ACGTGACAGAGACCCAGGACGAGAAATACCTACTCAATTCAGACGGAACCGGCAACT 487  
 DB 452 AAATGACAGAAACCCAGCTGAGATAAGAAATACCTACTTTTATTCAGAGTGAAGCTACCAAT 511  
 QY 488 ACAAGTACTGTTACAGTGAATGTAAGAGATACACAGCTGTATGTCTAAGGACACTT 547  
 DB 512 ACACAATATGTTTACAGTGAATGTAAGAAATACCTACTTTTATTCAGAGTGAAGCTACCAAT 571  
 QY 548 ACTTTAGGAAGATGAAACACAGGATGCACTCTCTGCAATCTCGAGGCTTTCGGAGC 607  
 DB 572 ACTTTAGAAAATGAAACACAGGACGCTCTGCTGCAATCTCGAGGCTTTCGGAGC 631  
 QY 608 CCACTGTGGAGTGGTGTCTGTCAGCTCCACAGGGAAGCTGTAAAGAAAGGCGCCTG 667

DB 632 CGATCGTGAATGGTGTCTTGGGATTCACAGGGGGAAGCTGTAAAGAAAGGCTCCAG 691  
 QY 668 CTGTTGTCAAGAAAGAGGAAAGGTACTTATGATGTTGTTCGGAACAGACATCAGATGCT 727  
 DB 692 CTGTTGTTAAAGAGGAGGAAAGGTCTTCAATGATTTATTTGGGCGGACATAAGGTGCT 751  
 QY 728 GTGCTAGAAATGCACTGGGCGCGGAATGCAACAGCTGTTTACCATAGATCTAAACCAGG 787  
 DB 752 GTGCCAGAAATGAACTGGGCGAGGGAATGCACAGGCTGTTTCAATAGATCTAAATCAAA 811  
 QY 788 CTCCTCAGAGCACACTGCCCGAGTTTATTCCTCAAGTGGGGAAACCTTGTGGATCAGGT 847  
 DB 812 CTCCTCAGACCACTGTCACAAATTTTCTTAAAGTAGGGGAACCTTATGGATAGGT 871  
 QY 848 GTAAGGCCATCCATGTGAACCAATGGAATCGGCTCACCTGGGAGCTGGGAAGCAAAAGCCC 907  
 DB 872 GCAAGCTGTTTATGTGAACCAATGGAATCGGCTCACCTGGGAAATTAGAAAACAAAGCAC 931  
 QY 908 TGGAGGAGGCGAGCTACTTTGAGATGAGTACCTACTCCNAAACAGGACCATGATTCGGA 967  
 DB 932 TCGAGGAGGCGCAACTTCTTGGATGAGTACCTATTTCAACAAACAGAACTATGATACGA 991  
 QY 968 TTCTCTGCGCTTGTGTCCTTCGTTGGGAAGAAACGACACCGGATATTACACTGCTCTT 1027  
 DB 992 TTCTGTTGCTTTGTATCATCATGTGGCAAGAAACGACACCGGATCTACACTTGTCTCT 1051  
 QY 1028 COTCAAAGCACCCCGAGCGTGTGGTGACCATCTCTAGAAAAGGGTTTATAAAGC 1087  
 DB 1052 CTTCAAAGCATCCCGAGTCAATCAGCTTTGGTTTACCATCGTAGAAAAGGATTTATAATG 1111  
 QY 1088 CTACAGCTCGCAGAAAGATGAAATTTGACCCCGTACGAAAGTTTCTGCTCTCAGTCA 1147  
 DB 1112 CTACCAATTTCAAGTGAAGATTTGAAATTTGACCAATATGAAGAGTTTCTGTTCTGCTCA 1171  
 QY 1148 GGTTTAAAGCGTACCCACGAATCCGATGCACTGATCTTCTCTCAGCCTCATTTCCCT 1207  
 DB 1172 GGTTTAAAGCGTACCCACGAATCAGATGATGAGTGTGAGCCTTCTCGAAATCATTTCCCT 1231  
 QY 1208 GTAAACAGAGAGGCTCGGAGGATGGGTACAGCATATCTAAATTTTGGCATCATAGAAAC 1267  
 DB 1232 GTAGCAAAAGGCTCTTGTATACGATACAGATATCCAGTTTGGCATCATAGACAC 1291  
 QY 1268 AGCCAGAGAGTACATATTTCTATGCAAAATGATGACGCCAGTTCACCAAAATGTTCA 1327  
 DB 1292 AGCCAGGAGATATATATTTCCATGCAAGAAATGATGATGCCCAAATTTACCAAAATGTTCA 1351  
 QY 1328 CGCTGAATATAGAAAGAAACCTCAAGTGTAGCAAAATGCTCAGCCAGCGAGGCTCT 1387  
 DB 1352 CGCTGAATATAGAAAGAAACCTCAAGTGTGCAAGAGCATCGGCAAGTCAAGGCTCTCT 1411  
 QY 1388 GTTCTCTGATGGCTTACCCGCTACCTCTTGGACCTGGAAGAGTGTTCGGACAAATCTC 1447  
 DB 1412 GTTCTCGATGGATACCAATTTCCATCTTGGACCTGGAAGAGTGTTCAGACAGTCTC 1471  
 QY 1448 CCAATTGCAAGGAGAAATCCCAAGAGGAGTTTGAATAAAAGGCTAACAGAAAGTGT 1507  
 DB 1472 CCACTGCAAGAGATACAGAAAGGAGTCTGGAATAGAAAGGCTAACAGAAAGTGT 1531  
 QY 1508 TTGGCCAGTGGGTGTGAGCAGTACTCTAAATATAGTGTAGGCGCGGGAAGGCTTCTGG 1567  
 DB 1532 TTGGACAGTGGGTGTGAGCAGTACTCTAAACATGAGTGAAGCCATAAAGGCTTCTGG 1591  
 QY 1568 TCBAATGCTGTGCGTACAAATTTCTATGGGCACTTGTGGAAACCATCTTTTAACTCAC 1627  
 DB 1592 TCAAGTGTGTGATACAAATTTCCCTTGGACATCTTGTGAGACGATCTTTTAACTCTC 1651  
 QY 1628 CAGGCCCTTCCCTTTTCAATCAAGACCAATCTCTCTTATCGGACCATTTGGGCTCTGTC 1687  
 DB 1652 CAGGCCCTTCCCTTTTCAATCAAGACCAATCTCTTATCTATGCAACAAATTTGGTGTTC 1711  
 QY 1688 TCCCTTCAATGTTGTTCTCATTTGTTGATCTGCCCAAAATACAAAAGCAATTTAGGT 1747

1712 TCCTCTTCATTGTCGTTTTAAACCTGCTAATTTGTCAAGTACAAAAAGCAATTTAGGT 1771  
1748 ACAGAGTCAGCTGCAGATGATCCAGGTGACTGGCCCCCTGGATACAGGACTACTCTACG 1807  
1772 ATGAAAGCAGCTACAGATGGTACAGGTGACCGGCTCTCAGATAATGACTTCTTACG 1831  
1808 TTGACTTCAGGGACTATGAATATGACCTTAAGTGGGAGTTCCCGAGAGAGAACTTAGAGT 1867  
1832 TTGATTTTCAGAGATATGATATGATCTCAATGGGAGTTTCCAGAGAGAAATTTAGAGT 1891  
1868 TTGGGAAGGTCCTGGGGTCTGGCGCTTTCCGGAGGGTGTAAAGCGCCACGGCCCTATGGCA 1927  
1892 TTGGGAAGGTCATAGGATCAGGTGCTTTTGGAAAAGTGTAAAGCGCAACAGCTTATGGAA 1951  
1928 TTAGTAAACCGGAGTCTCAATTCAGGTGGCGGTGAAGTGTCTAAAGAGAGAAAGCTGACA 1987  
1952 TTAGCAAAACAGGAGTCTCAATTCAGGTGGCGGTCAAAATGTCTGAAAGAAAAGCAGACA 2011  
1988 GCTGTGAAAAAGAGCTCTCATGTCGGAGTCAAAATGATGACCCACCTGGGACACCATG 2047  
2012 GCTCTGAAAGAGGACCTCACTGTCAGACTCAAGATGATGACCCAGCTGGGAGCCACG 2071  
2048 ACAACATCGTGAATCTGTGGGGCATGCACACTGTCCAGGGCCAGTGTACTTGAATTTTG 2107  
2072 AGAATATTGTGAACCTGTGGGGCGTGCACACTGTCCAGGACCAATTTACTTGAATTTTG 2131  
2108 AATATTGTGCTATGGTGACCTCTCACTAATCTAAAGAGTAAAGAGAGAAAGTTTCA 2167  
2132 AATACTGTGTGATGGTGATCTCTCAACTATCAAGAGTAAAGAGAGAAATTTCA 2191  
2168 GGACATGGACAGAGATTTTAAAGGAACATAATTCAGTTTCACTACCCGCTTGGATCAGC 2227  
2192 GGACATTGGACAGATTTTCAAGGAACACAATTTTCACTTACCCCACTTTCCCAATCAC 2251  
2228 ATTCAAAATCCAGCATGCTGTTTACAGAGAGTTCAAGTTACACCCGCTTGGATCAGC 2287  
2252 ATCCAAATTTCCAGCATGCTGTTTCAAGAGAGTTCAAGTACACCCGAGTCCGATCAA 2311  
2288 TCTCAGGTTCAATGGGAATTCATTCATCTGAAGATGAGATGGAATGAAACCCAGA 2347  
2312 TCTCAGGGCTTCATGGGAATTCATTTCACTGGAAGTGAATGAAATGAAACCCAAA 2371  
2348 AGAGCTGGCAGAGAGAGAGGAGAGATTTCAAGTGTGAGTGTGAGACCTTCTTT 2407  
2372 AAAGGCTG-----GAAAGAGAGGAGGACTGAATGTGCTTACATTTGAAGATCTTCTTT 2425  
2408 GCTTTGCGTACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAAGTGTGTCCTCCACA 2467  
2426 GCTTTGCAATCAAGTGGCCAAAGGATGGAATTTCTGGAATTTAAGTGTGTTTCA 2485  
2468 GAGACTGGCAGCCAGGAATGTGTTGGTCAACCCAGGAAGTGTGAGATCTGTGACT 2527  
2486 GAGACTGGCCGCCAGGAACGTTGTTGTCACCAAGGAAGTGTGAGATGATGACT 2545  
2528 TTGAGCTGGCCGAGACATCTGAGCGACTCCAGCTACGTCTGAGGGGCAAGCAGCGC 2587  
2546 TTGGATTGGCTCGAGATATCATGAGTGATTCCAACTATGTGTCTGAGGGCAATGCCCGTC 2605  
2588 TGCCGCTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGATCTACACAATCAAGAGTG 2647  
2606 TGCCGTGAATGATGGCCCGCCGAAAGCTGTTTGAAGGATCTACACCAATTAAGATG 2665  
2648 AGCTGTGCTTACCGCATCTCTCTGGAGATATTTTCACTGGGTGTGAACCTTACC 2707  
2666 ATGTCTGTCTATGGAATATTTACTGTGGGAATCTTCTCACTGTGTGATCTTACC 2725  
2708 CTGGCATTCCTGTGACCGCTAACTTCTATAAACTGATTCAGAGTGGAATTTAAATGGAGC 2767  
2726 CTGGCATTCCTGTGATGCTAACTTCTACAACTGATTCAAAATGGATTTAAATGGATC 2785  
2768 AGCCATTCATGCCACAGAGGATATCTTTGTAATGCAATCCCTGCTGGGCTTTTGACT 2827  
2786 AGCCATTTTATGCTACAGAGAAATATACATTAATGAATTCCTGCTGGGCTTTTGACT 2845

2828 CAAGGAAGCGGCGCATCTCTCCCAACCTGACTTATTTTAGGATGTCACTGGCAGAGG 2887  
2846 CAAGGAAGCGGCGCATCTCTCCCAATTTGACTTCTGTTTTTAGGATGTCACTGGCAGATG 2905  
2888 CAGAGCAAG-----CATGTATCAGAAACATCCATCTCATCTACC 2924  
2906 CAGAGCAAGCGATGTATCAGAAATGTGGATGGCGCTGTTTCGGAATGTCTCTCACACCTACC 2965  
2925 AAAACAGGCGGCGCCCTCAGCAGAGAGCGGGCTCAGAGC--CCAGTCCGCCACAGCGCCAGG 2983  
2966 AAAACAGGCGACCTTTCAGCAGAGATGGAATTTGGGGTACTCTCTCCGAGGCTCAGG 3025  
2984 T-GAAGATTTCAGAGAAAGAAAGTTAGCAGAGAGGCTTTGGACCCCGCC-----ACCTAG 3038  
3026 TCGAAGATTTCGTAGAGGAACAATTTAGTTTAAAGGACTTCATCCCTCCACCTATCCCTAA 3085  
3039 CAGGCTGTAGACCGCAGAGCCCAAGATTAGCTCGCTCT--GAGGAAGCGCCCTACAGCG 3096  
3086 CAGGCTGTAGATTACAAAACAAGATTAAATTTCACTCACTAAAAGAAAATCTATTATCAAC 3145  
3097 CGTTGCTTCGCTGGACTTTTCTCTAGATGCTGTCTGCCATTACTC-----CAAAAGTGA 3149  
3146 TGCTGCTTCACAGACTTTTCTCTAGAGCCGTCTGGCTTTACTCTTGTTTTCAAAAGGGA 3205  
3150 CTTCTATAAAATCAAACTCTCTCTCCACAGCGGAGAGCCAAATAATGAGACTTGTGG 3209  
3206 CTTTTGTAATAATCAAAATCATCTGTCAAGGCGAGGAGCTGATAATGAATCTTTATTG 3265  
3210 TGAGCCGCGCTACCTGGGGCTTTCCACAGAGCTTTGAGGGAAAGCCATGTATCTGAAA 3269  
3266 GAGCAATTGATGTCATCCAAAGCCCTTCTCAGCGCGCTTGTAGTGAATTTGTACTCTGAAG 3325  
3270 TATAGTATATTCTGTAAATACGTCAAAACAAACCAACCCGTTTTTTGCTAAGGGAAGC 3329  
3326 TACAGTATATTCTGTAAATACATAAAACAA-----AGCATTTTGTAAAGGAGAGC 3378  
3330 TAAATATGATTTTTAAAAATCTATGTTTAAATACTATGTAACCTTTTTTCATCTATTAG 3389  
3379 TAATATGATTTTT--AAGTCTATGTTTAAATAATATGTAATTTTTCAGCTATTAG 3436  
3390 TGATATTTTTATGATGGAATAAACTTTCTACTGTAAAAAAAAAATAAAAAA 3449  
3437 TGATATTTTTATGGTGGGAATAAAATTTCTACTACAGAAAAAAAATAAAAAA 3496  
3450 AAAA 3453  
3497 AAAA 3500

RESULT 15  
AAT38734  
ID AAT38734 standard; cDNA; 3501 BP.  
XX AAT38734;  
AC AAT38734;  
XX 25-MAR-2003 (updated)  
DT 11-DEC-1996 (first entry)  
XX Murine foetal liver kinase 2 cDNA.  
XX  
KW Murine; foetal liver kinase 2; flk-2; protein tyrosine kinase;  
KW monoclonal; antibody; extracellular domain; receptor assay;  
KW haematopoietic stem cell; ligand; stimulation; proliferation;  
KW differentiation; treatment; anaemia; bone marrow damage;  
KW cancer chemotherapy; radiation; ds.  
XX Mus musculus.  
OS  
XX  
FH Key Location/Qualifiers  
FT CDS 58..3039  
FT   
FT sig\_peptide /\*tag= a 58..138



QY 1448 CCAATTGCA CGGAGGAAATCCAGAAAGAGTTTGAATATAAAGGCTAAACAGAAAAGTGT 1507  
DB 1472 CCAACTGCA CAGAGAGATCACAGAGAGTCTGGAATAGAAAGCTAAACAGAAAAGTGT 1531  
QY 1508 TTGGCCAGTGGGTGTGCGAGCAGTACTCTAAATATGAGTGAAGCGCGGAAAGGGCTTCTGG 1567  
DB 1532 TTGGAAGTGGGTGTGCGAGCAGTACTCTAAACATGAGTGAAGCCATAAAAGGGTTCCTGG 1591  
QY 1568 TCAAAATGCTGCGGTACAAATCTATGGGCGAGTCTTGGGAAACCATCTTTTAAACTCAC 1627  
DB 1592 TCAAGTGTCTGCGATACAAATCTTCCCTTGGGCATCTTGTGAGCGATCTTTTAAACTCTC 1651  
QY 1628 CAGGCCCTTCCCTTTTCATCCAGACAAACATCTCTCTATGCGACCATTTGGGCTGTGC 1687  
DB 1652 CAGGCCCTTCCCTTTTCATCCAGACAAACATCTCTATGCGAAACATTTGGTGTGTC 1711  
QY 1688 TCCCTTCATTTGTTGTTCTCATTTGTTGATGTCGCCAAATAACAAAAGCAATTTAGT 1747  
DB 1712 TCCCTTCATTTGTTGTTTAAACCTGCTAATTTGTCAAGATCAAAAAGCAATTTAGT 1771  
QY 1748 ACGAGATGAGTGCAGATGATCAGGTGACTGGCCCCCTGGATACAGAGTACTTCTACG 1807  
DB 1772 ATGAAAGCCAGCTACAGATGATGAGTGACCGGCTCTCAGATATAGTACTTCTACG 1831  
QY 1808 TTGACTTCAGGGACTATGAATATGACCTTAAGTGGGAGTTCCCGAGAGAGAACTTAGGT 1867  
DB 1832 TTGATTTAGAGAAATATGAATATGATCTCAATGGGAGTTTCAAGAGAAAATTTAGGT 1891  
QY 1868 TTGGGAAGTCTCTGGGCTCTGGCGCTTTCCGGAGGGTGATGAACGCCACCGCCTATGGCA 1927  
DB 1892 TTGGGAAGTCTAGGATCAGGTGCTTTTGGAAAGTGATGAACGCAACAGCTTATGGAA 1951  
QY 1928 TTAGTAAACCGGAGTCTCAATTCAGTGGCGGTGAAGATGCTTAAAGAGAAAGCTGACA 1987  
DB 1952 TTAGCAAAAACAGAGTCTCAATCCAGGTGCGCTCAAAATGCTGAAAGAAAAGCAGACA 2011  
QY 1988 GCTGTCAAAAAGAGCTCTCATGCTCGGAGCTCAAAATGATGACCCACCTGGGACCATG 2047  
DB 2012 GCTGTGAAAGAGAGGCACTCATGTGAGAACTCAAGATGATGACCCAGCTGGGAGCCAG 2071  
QY 2048 ACAACATCGTGAATCTGTGGGGGCAATGCACACTGTGAGGGCCAGTGTACTGATTTTG 2107  
DB 2072 AGAATATTGTGAACCTGTGGGGGCGTGACACTGTGAGGACCAATTTACTGATTTTG 2131  
QY 2108 AATATTGTTGCTATGCTGACCTCTCAACTAAGAAAGTAAAGAGAGAAATTTTACA 2167  
DB 2132 AATACTGTTGCTATGCTGATCTCTCAACTAAGAAAGTAAAGAGAGAAATTTTACA 2191  
QY 2168 GGACATGGACAGAGATTTTAAAGAACATAATTTGATCTTACCTACTTTCCAGGCAC 2227  
DB 2192 GGACTTGGACAGAGATTTTCAAGGAAACAAATTTCAAGTTTACCCCACTTTCCATCAC 2251  
QY 2228 ATTCAAATTCAGACGCTGCTGTTCAGAGAGTTCAGTTTACACCGCCCTTGGATCAGC 2287  
DB 2252 ATCCAAATTCAGATGCTGCTGTTCAGAGAGTTCAGATACACCGGACTCGATCAAA 2311  
QY 2288 TCTCAGGGTTCATAGGAAATCAATTCATCTGAAGATGAGATGGAATGAAAACAGA 2347  
DB 2312 TCTCAGGGTTCATAGGAAATCAATTCATCTGAAGATGAAATGGAATGAAAACAGA 2371  
QY 2348 AGAGGCTGSCAAGAGAGAGGAGAGATTTGAACTGCTGAGTTTGAAGACCTCTTT 2407  
DB 2372 AAAGGCTG-----GAAGAGAGAGGAGGACTTGAATGTGCTTACATTTGAAGATCTCTTT 2425  
QY 2408 GCTTTGCGTACCAAGTGGCCAAAGGATGGAATTCCTGAGTTCAAGTCTGTGTCACA 2467  
DB 2426 GCTTTGCTATCAAGTGTGCCAAGGAATGGAATTTCTGGAATTTAAGTCTGTGTTACA 2485  
QY 2468 GAGACTGCGACCCAGGAATGTGTTGGTCAACCCAGGAGGTGGTGAAGATCTGTGACT 2527  
DB 2486 GAGACTGCGCCAGGAACGTGCTTGTCAACCCAGGGAAGTGGTGAAGATATGTGACT 2545  
QY 2528 TTGACTGGCCCGAGACATCTGTAGGACTCCAGCTACGTCAGGGGCAACGCAACGGC 2587

DB 2546 TTGGAATGGCTCGAGATATCATGAGTGAATTCGAATGTTGTGAGGGGCAATGCCGTC 2605  
QY 2588 TGCCGGTGAAGTGGATGGCAACCCGAGAGCTTATTTGAAGGATCTACACAATCAAGAGTG 2647  
DB 2606 TGCCGTGAATATGATGGGCCCCCGAAGCCCTGTTGAAGGCACTACACCAATTAAGAGTG 2665  
QY 2648 ACCTCTGGTCTTACGGCATCTTCTCTGGGAGATATTTTCACTGGGTGTGAACCTTACC 2707  
DB 2666 ATGCTGGTCAATGGAATATTAATCTGTGGGAAATCTTCTCACTGGTGTGAATCCTTACC 2725  
QY 2708 CTGSCATTCCTGTCGAGCTAACTCTATFAAATGATTTCAAGTGAATTTAAATGAGC 2767  
DB 2726 CTGSCATTCCTGTCGAGCTAACTCTTCAAACTGATTTCAAAATGGAATTTAAATGAGC 2785  
QY 2768 AGCCATTCATGTCACAGAGGATATCTTGTATGCAATCTCTGGGCTTTTTCGACT 2827  
DB 2786 AGCCATTTATGCTACAGAGAAATATACATTAATATGCAATCTCTGGGCTTTTTCGACT 2845  
QY 2828 CAAGGAAGCGGCGCATCTTCCCAACCTGACTTCAATTTTATGATGTCAGCTGGCAGAG 2887  
DB 2846 CAAGGAAGCGGCGCATCTTCCCTTAATTTGACTTCGTTTATGAGTGTGAGTGGCAGATG 2905  
QY 2888 CAGAGAAG-----CATGTATCAGAAACATCCATTCATCTACC 2924  
DB 2906 CAGAGAAGCGATGTATCAGAAATGTGGATGGCGCTGTTCCGAATGTCTTCAACCTACC 2965  
QY 2925 AAAACAGGCGGCGCTCAGCAGAGAGCGGCTCAGAGC--CAGTCCGCCACAGGCCAGG 2983  
DB 2966 AAAACAGGCGACCTTTCAGCAGAGAGATGATTTGGGCTACTCTCTCCGCGGCTCAGG 3025  
QY 2984 T-GAAGATTCACAGAGAAAGATTTAGCAGAGAGGCTTTGGACCCCGCC-----ACCCTAG 3038  
DB 3026 TCGAAGATTCGTAGAGGAACAAATTTAGTTTAAAGGACTTCATCCCTCCACTATCCCTAA 3085  
QY 3039 CAGGCTGTAGACCGCAGAGCCGAAGATTTAGCTCGCTCT--GAGGAAGGCCCTTACAGCG 3096  
DB 3086 CAGGCTGTAGATTTACCAAAACAGATTAATTTTCACTAAAGAAATCTATTTATCAAC 3145  
QY 3097 CGTGTCTCGCTGGACTTTTCTAGATGCTGTCTGCCATTACTC-----CAAAGTGA 3149  
DB 3146 TGTGCTTCCACAGACTTTTCTAGAGCGCTCTCGTTTACTCTTGTCTTCAAAAGGGA 3205  
QY 3150 CTTCTATAAATCAAACTCTCTCTCGCAGAGCGGAGAGCCAAATATGAGACTTGTGG 3209  
DB 3206 CTTTGTAAATCAAACTCATCTGTGTCAGAGGAGGAGCTGATATGAACCTTTATTG 3265  
QY 3210 TGAGCCCGCTTACCTGGGGCTTTTCCAGAGCTTTGAGGGGAAAGCCATGTATCTGAAA 3269  
DB 3266 GAGCATTTGATCTGCATCCAAAGGCTTCTCAGGCGGCTTGAAGTGAATTTGTACTGAG 3325  
QY 3270 TATAGTATATCTTCTTAATACGTGAAACAAACCAACCCGTTTTCCTAGGGGAAAGC 3329  
DB 3326 TACAGTATATCTTGTAAATACATAAAACAAA-----AGCATTTTGTAGGAGAAAGC 3378  
QY 3330 TAAATATGATTTTAAATAATCTATGTTTAAATACTATGTAATCTTTTTCATCTATTAG 3389  
DB 3379 TAAATATGATTTT--AGTCTATGTTTAAATAATATGTAATTTTTCAGCTATTAG 3436  
QY 3390 TGATATATTTATGATGGAATAAACTTTTCTACTGTAAAAAATAAATAAATAAATAAATAA 3449  
DB 3437 TGATATATTTATGGTGGGAATAAAATTTCTACTACAGAAAAAATAAATAAATAAATAA 3496  
QY 3450 AAAA 3453  
DB 3497 AAAA 3500

Search completed: August 27, 2003, 16:01:32  
Job time : 1042.85 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw.model

Run on: August 28, 2003, 01:02:19 ; Search time 1007.5 Seconds  
(without alignments)  
7863.095 Million cell updates/sec

Title: US-09-919-408A-1  
Perfect score: 3453  
Sequence: 1 GCGGCTGGTACCGCGCGC.....AAAAAAAAAAAAAAAAAAAA 3453

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 1533700 seqs, 1147125425 residues

Total number of hits satisfying chosen parameters: 3067400

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/1/pubna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/1/pubna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/1/pubna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/1/pubna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/1/pubna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/1/pubna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/1/pubna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/1/pubna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/1/pubna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/1/pubna/US10A\_PUBCOMB.seq:  
14: /cgn2\_6/ptodata/1/pubna/US10B\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/1/pubna/US10\_NEW\_PUB.seq:  
16: /cgn2\_6/ptodata/1/pubna/US60\_NEW\_PUB.seq:  
17: /cgn2\_6/ptodata/1/pubna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	3453	100.0	3453	9	US-09-919-408-1	Sequence 1, Appli
2	3453	100.0	3453	10	US-09-872-136-1	Sequence 1, Appli
3	2227.2	64.5	3501	9	US-09-919-408-3	Sequence 3, Appli
4	2227.2	64.5	3501	10	US-09-872-136-3	Sequence 22, Appli
5	2081.8	60.3	3120	10	US-09-982-610-22	Sequence 7538, Ap
6	466.6	13.5	645	10	US-09-796-692-7538	Sequence 7538, Ap
7	466.6	13.5	645	14	US-10-040-862-7538	Sequence 8085, Ap
8	456.4	13.2	610	10	US-09-796-692-8085	Sequence 8085, Ap
9	456.4	13.2	610	14	US-10-040-862-8085	Sequence 8927, Ap
10	454.6	13.2	605	10	US-09-796-692-8927	Sequence 8927, Ap
11	454.6	13.2	605	14	US-10-040-862-8927	Sequence 7976, Ap
12	448.8	13.0	597	10	US-09-796-692-7976	Sequence 7976, Ap
13	448.8	13.0	597	14	US-10-040-862-7976	Sequence 8273, Ap
14	448	13.0	601	10	US-09-796-692-8273	Sequence 8273, Ap
15	448	13.0	601	14	US-10-040-862-8273	Sequence 8132, Ap
16	448	13.0	608	10	US-09-796-692-8132	

C	17	448	13.0	608	14	US-10-040-862-8132	Sequence 8132, Ap
	18	447.2	13.0	597	10	US-09-796-692-9010	Sequence 9010, Ap
	19	447.2	13.0	597	14	US-10-040-862-9010	Sequence 9010, Ap
	20	440	12.7	605	10	US-09-796-692-8047	Sequence 8047, Ap
	21	440	12.7	605	14	US-10-040-862-8047	Sequence 8047, Ap
	22	439.8	12.7	589	10	US-09-796-692-8630	Sequence 8630, Ap
	23	439.8	12.7	589	14	US-10-040-862-8630	Sequence 8630, Ap
	24	424.4	12.3	594	10	US-09-796-692-8135	Sequence 8135, Ap
	25	424.4	12.3	594	14	US-10-040-862-8135	Sequence 8135, Ap
	26	423.2	12.3	564	10	US-09-796-692-9025	Sequence 9025, Ap
	27	423.2	12.3	564	14	US-10-040-862-9025	Sequence 9025, Ap
	28	416.8	12.1	556	10	US-09-796-692-7883	Sequence 7883, Ap
	29	416.8	12.1	556	14	US-10-040-862-7883	Sequence 7883, Ap
	30	407.2	11.8	550	10	US-09-796-692-8494	Sequence 8494, Ap
	31	407.2	11.8	550	14	US-10-040-862-8494	Sequence 8494, Ap
C	32	399.8	11.6	539	10	US-09-796-692-8255	Sequence 8255, Ap
C	33	399.8	11.6	539	14	US-10-040-862-8255	Sequence 8255, Ap
	34	386.8	11.2	510	10	US-09-796-692-8268	Sequence 8268, Ap
	35	386.8	11.2	510	14	US-10-040-862-8268	Sequence 8268, Ap
	36	360.2	10.4	497	10	US-09-796-692-8408	Sequence 8408, Ap
	37	360.2	10.4	497	14	US-10-040-862-8408	Sequence 8408, Ap
	38	357.4	10.4	3992	10	US-09-944-807-9	Sequence 9, Appli
	39	357.4	10.4	3992	12	US-10-007-326A-292	Sequence 292, App
	40	357.4	10.4	4008	12	US-10-101-510-544	Sequence 544, App
	41	350	10.1	606	10	US-09-796-692-8114	Sequence 8114, Ap
	42	350	10.1	606	14	US-10-040-862-8114	Sequence 8114, Ap
	43	342.2	9.9	508	10	US-09-796-692-8032	Sequence 8032, Ap
	44	342.2	9.9	508	14	US-10-040-862-8032	Sequence 8032, Ap
	45	329	9.5	5084	10	US-09-967-768A-277	Sequence 277, App

ALIGNMENTS

RESULT 1  
US-09-919-408-1  
; Sequence 1, Application US/09919408  
; Patent No. US20020072077A1  
; GENERAL INFORMATION:  
; APPLICANT: Lemischke, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ImClone Systems Incorporated  
; STREET: 180 Varick Street  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/09/919,408  
; FILING DATE: 31-Jul-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/977,451  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: US 07/906,397  
; FILING DATE: 26-JUN-1992  
; APPLICATION NUMBER: US PCT/US92/05401  
; FILING DATE: 26-JUN-1992  
; APPLICATION NUMBER: TW 81102961  
; FILING DATE: 15-APR-1992  
; APPLICATION NUMBER: US PCT/US92/02750  
; FILING DATE: 02-APR-1992  
; APPLICATION NUMBER: US 07/813,593  
; FILING DATE: 24-DEC-1991  
; APPLICATION NUMBER: US 07/793,065



1621 AACTCACCAGGCCCTTCCCTTTCATCCAGACAAACATCTCTTCTATGCGACCATGGG 1680  
1621 AACTCACCAGGCCCTTCCCTTTCATCCAGACAAACATCTCTTCTATGCGACCATGGG 1680  
1681 CTCTGCTCCCTTTCATGTTGTTCTCATTTGTTGATCTGCCACAATAACAAAAGCAA 1740  
1681 CTCTGCTCCCTTTCATGTTGTTCTCATTTGTTGATCTGCCACAATAACAAAAGCAA 1740  
1741 TTTAGGTACGAGAGTCAAGTGCAGATGCCAGGTGATCGGCCCTTGGATACAGGATAC 1800  
1741 TTTAGGTACGAGAGTCAAGTGCAGATGCCAGGTGATCGGCCCTTGGATACAGGATAC 1800  
1801 TTCTACGTTGACTTCAGGAGTATGAATATGACCTTAAGTGGGAGTTCCTCCGAGAGAAC 1860  
1801 TTCTACGTTGACTTCAGGAGTATGAATATGACCTTAAGTGGGAGTTCCTCCGAGAGAAC 1860  
1861 TTAGAGTTTGGAGAGTCTGGGGTCTGGCGCTTTCGGGAGGTTGATGAAGCCACGGCC 1920  
1861 TTAGAGTTTGGAGAGTCTGGGGTCTGGCGCTTTCGGGAGGTTGATGAAGCCACGGCC 1920  
1921 TATGGCAATTAGTAAACCGGAGTCTCAATTCAAGTGGCGGTGAAGATGCTTAAAGAGAAA 1980  
1921 TATGGCAATTAGTAAACCGGAGTCTCAATTCAAGTGGCGGTGAAGATGCTTAAAGAGAAA 1980  
1981 GCTGACAGCTGTGAAAAGAGTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA 2040  
1981 GCTGACAGCTGTGAAAAGAGTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA 2040  
2041 CACATGACAACTCGTGAATCTCGTGGGCGATGCACACTGTCAGGGCCAGTGTACTTG 2100  
2041 CACATGACAACTCGTGAATCTCGTGGGCGATGCACACTGTCAGGGCCAGTGTACTTG 2100  
2101 ATTTTGAATATTGTTGCTGATGCTCACTCACTCACTCACTCACTCACTCACTCACTCA 2160  
2101 ATTTTGAATATTGTTGCTGATGCTCACTCACTCACTCACTCACTCACTCACTCACTCA 2160  
2161 TTTTCAAGGACATGACAGAGATTTTAAAGAAATATTTTCAAGTGTCTTACCTACTTTC 2220  
2161 TTTTCAAGGACATGACAGAGATTTTAAAGAAATATTTTCAAGTGTCTTACCTACTTTC 2220  
2221 CAGGACATTTCAAAATTCAGATGCTTGGTTCACGAGAGTTCAGTTACCCCGCCCTTG 2280  
2221 CAGGACATTTCAAAATTCAGATGCTTGGTTCACGAGAGTTCAGTTACCCCGCCCTTG 2280  
2281 GATCAGCTCTCAGGTTCAATGGGAATTCATCTGAGAGATGAGATTAATATGAA 2340  
2281 GATCAGCTCTCAGGTTCAATGGGAATTCATCTGAGAGATGAGATTAATATGAA 2340  
2341 AACAGAGAGGCTGGCAGAGAAAGAGGAGGAGATTTTGAACGCTGACCTTTGAGAC 2400  
2341 AACAGAGAGGCTGGCAGAGAAAGAGGAGGAGATTTTGAACGCTGACCTTTGAGAC 2400  
2401 CTCCTTTGCTTGGTACCAAGTGGCCAAAGCATGGAATTCCTGGAGTTCAGTCTGTT 2460  
2401 CTCCTTTGCTTGGTACCAAGTGGCCAAAGCATGGAATTCCTGGAGTTCAGTCTGTT 2460  
2461 GTCCACAGAGACCTGGCAGCAGAGATGTTGTTGTCACCCAGGAGGTGTGAAGATC 2520  
2461 GTCCACAGAGACCTGGCAGCAGAGATGTTGTTGTCACCCAGGAGGTGTGAAGATC 2520  
2521 TGTGACTTTGGACTGGCCGAGACATCTGAGCGACTCCAGCTAGCTCGTCAGGGGCAAC 2580  
2521 TGTGACTTTGGACTGGCCGAGACATCTGAGCGACTCCAGCTAGCTCGTCAGGGGCAAC 2580  
2581 GCACGGCTGCCGTTGAGTGAATGACCGAGAGCTTATTTGAAGGATCTACCAATC 2640  
2581 GCACGGCTGCCGTTGAGTGAATGACCGAGAGCTTATTTGAAGGATCTACCAATC 2640  
2641 AAGAGTGAAGTCTGCTCAGGATCTCTCTGGGAGATTTTTCAGTGGTGTGAAC 2700  
2641 AAGAGTGAAGTCTGCTCAGGATCTCTCTGGGAGATTTTTCAGTGGTGTGAAC 2700  
2701 CCTTACCCTGGCAATCTCTGTCGACGCTAACTTCTATATAAACTGATTCAGAGTGGATTTAA 2760

2701 CCTTACCCTGGCAATCTCTGTCGAGCTTAATCTTCTATATAAATGATTCAGAGTGGATTTAA 2760  
2761 ATGAGCAGAGCCATTCTATGCCACAGAGGAGATATACCTTTGTAATGCAATCTCTGCTGGGCT 2820  
2761 ATGAGCAGAGCCATTCTATGCCACAGAGGAGATATACCTTTGTAATGCAATCTCTGCTGGGCT 2820  
2821 TTTGACTCAAGGAAGCGGCCATCTCTCCCAACCTGACTTCTATTTTAAAGATGTCAGCTG 2880  
2821 TTTGACTCAAGGAAGCGGCCATCTCTCCCAACCTGACTTCTATTTTAAAGATGTCAGCTG 2880  
2881 GCAGAGCAGAGGAAGCATGTATCAGAAACATCCATCTCATCTACCAAAACAGCGGCCCT 2940  
2881 GCAGAGCAGAGGAAGCATGTATCAGAAACATCCATCTCATCTACCAAAACAGCGGCCCT 2940  
2941 CAGCAGAGAGCGGGCTCAGAGCCAGTCGTCACAGCGCCAGGTGAAGATTCACAGAGAA 3000  
2941 CAGCAGAGAGCGGGCTCAGAGCCAGTCGTCACAGCGCCAGGTGAAGATTCACAGAGAA 3000  
3001 AGAAGTTAGCGAGAGCGCTTGGACCCCGCCACCTAGCAGGCTGTAGACCGCAGAGCA 3060  
3001 AGAAGTTAGCGAGAGCGCTTGGACCCCGCCACCTAGCAGGCTGTAGACCGCAGAGCA 3060  
3061 AGATTAGCTCGCTCTGAGGAGCGCCCTACAGCGGCTTCTCGCTGGACTTTTCTCT 3120  
3061 AGATTAGCTCGCTCTGAGGAGCGCCCTACAGCGGCTTCTCGCTGGACTTTTCTCT 3120  
3121 AGATGCTGTCTGCCATTACTCCAAAGTGACTTCTATAAAATCAAACTCTCTCGCACAG 3180  
3121 AGATGCTGTCTGCCATTACTCCAAAGTGACTTCTATAAAATCAAACTCTCTCGCACAG 3180  
3181 GCGGAGAGCCAAATATAGACTTTGTTGGTGAGCGCGCTACCTCGGGGGCTTTCCAGG 3240  
3181 GCGGAGAGCCAAATATAGACTTTGTTGGTGAGCGCGCTACCTCGGGGGCTTTCCAGG 3240  
3241 AGCTTGAGGGAGAACCATGTATCTGAAATATAGTATATTTTAAATCTATGTTTAA 3300  
3241 AGCTTGAGGGAGAACCATGTATCTGAAATATAGTATATTTTAAATCTATGTTTAA 3300  
3301 ACCAAACCCGTTTTTCTAGGGAAGCTAAATATGATATTTTAAATCTATGTTTAA 3360  
3301 ACCAAACCCGTTTTTCTAGGGAAGCTAAATATGATATTTTAAATCTATGTTTAA 3360  
3361 AATACTATGTAACCTTTTCTATCTATTTAGTGATATTTTATGATGGAATAAACTTTC 3420  
3361 AATACTATGTAACCTTTTCTATCTATTTAGTGATATTTTATGATGGAATAAACTTTC 3420  
3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAA 3453  
3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAA 3453

## RESULT 2

US-09-872-136-1  
; Sequence 1, Application US/09872136  
; Patent No. US20020119545A1  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Imclone Systems Incorporated  
; STREET: 180 Varick Street  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/872,136  
 FILING DATE: 01-Jun-2001  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/09/208,786  
 FILING DATE: <Unknown>  
 APPLICATION NUMBER: US/09/021,324  
 FILING DATE: <Unknown>  
 APPLICATION NUMBER: US/07/977,451  
 FILING DATE: 1992-11-19  
 APPLICATION NUMBER: US 07/906,397  
 FILING DATE: 26-JUN-1992  
 APPLICATION NUMBER: US PCT/US92/05401  
 FILING DATE: 26-JUN-1992  
 APPLICATION NUMBER: TW 81102961  
 FILING DATE: 15-APR-1992  
 APPLICATION NUMBER: US PCT/US92/02750  
 FILING DATE: 02-APR-1992  
 APPLICATION NUMBER: US 07/813,593  
 FILING DATE: 24-DEC-1991  
 APPLICATION NUMBER: US 07/793,065  
 FILING DATE: 15-NOV-1991  
 APPLICATION NUMBER: US 07/728,913  
 FILING DATE: 28-JUN-1991  
 APPLICATION NUMBER: US 07/679,666  
 FILING DATE: 02-APR-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Feit, Irving N.  
 REGISTRATION NUMBER: 28,601  
 REFERENCE/DOCKET NUMBER: LEM-3-7P  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 212-645-1405  
 TELEFAX: 212-645-2054  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3453 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: gDNA  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 FRAGMENT TYPE: N-terminal  
 FEATURE:  
 NAME/KEY: mat\_peptide  
 LOCATION: 112..3006  
 FEATURE:  
 NAME/KEY: sig\_peptide  
 LOCATION: 31..111  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 31..3009  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:

Query Match 100.0%; Score 3453; DB 10; Length 3453;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

181	CCATCATCGTACCGAATGGTGGAGGATCCCCAGAAAGACCTCCAGTGTACCCCGAGCGC	240
241	CAGAGTGAAGGACCGGTATATGAAGCGGCCACCGTGGAGGTGGCCGAGTCTGGGTCCATC	300
241	CAGAGTGAAGGACCGGTATATGAAGCGGCCACCGTGGAGGTGGCCGAGTCTGGGTCCATC	300
301	ACCCTGCAAGTGCAGCTCGCCA CCCAGGGGACCTTTCTCTGCTCTGGGTCTTTTAAGCAC	360
301	ACCCTGCAAGTGCAGCTCGCCA CCCAGGGGACCTTTCTCTGCTCTGGGTCTTTTAAGCAC	360
361	AGCTCCCTGGGCTGCCAGCGGCACCTTTGATTTACAAAACAGAGGAATCGTTTCCATGGCC	420
361	AGCTCCCTGGGCTGCCAGCGGCACCTTTGATTTACAAAACAGAGGAATCGTTTCCATGGCC	420
421	ATCTTGAACGTGCACAGAGACCCAGCAGAGAAATACCTACTCCATATTCAGAGCGAAGCG	480
421	ATCTTGAACGTGCACAGAGACCCAGCAGAGAAATACCTACTCCATATTCAGAGCGAAGCG	480
481	GCCAACTACACAGTACTGTTTCA CAGTGAATGTAAGAGATACACAGCTGTATGTCTAAGG	540
481	GCCAACTACACAGTACTGTTTCA CAGTGAATGTAAGAGATACACAGCTGTATGTCTAAGG	540
541	AGACCTTACTTTAGGAAGATGGAAAAACAGGATGCATGCTCTGCATCTCCGAGGGTGT	600
541	AGACCTTACTTTAGGAAGATGGAAAAACAGGATGCATGCTCTGCATCTCCGAGGGTGT	600
601	CGGAGCCCACTGTGGAGTGGGTGCTCTGCAGCTCCCA CAGGAAAGCTGTAAAGAGAA	660
601	CGGAGCCCACTGTGGAGTGGGTGCTCTGCAGCTCCCA CAGGAAAGCTGTAAAGAGAA	660
661	GGCCCTGCTGTTGT CAGAAAAGGAGGAAAAGGTACTTCA TGTGTTTCGAAACAGACATC	720
661	GGCCCTGCTGTTGT CAGAAAAGGAGGAAAAGGTACTTCA TGTGTTTCGAAACAGACATC	720
721	AGATGCTGTCTAGAAATGCACTGGGCGCGGAATGCACCAAGCTGTTCCATATATCTA	780
721	AGATGCTGTCTAGAAATGCACTGGGCGCGGAATGCACCAAGCTGTTCCATATATCTA	780
781	AACAGGGTCTCTCAGAGCACACTGCCAGTATTTCTCTGAAAGTGGGGGAAACCTTTGGG	840
781	AACAGGGTCTCTCAGAGCACACTGCCAGTATTTCTCTGAAAGTGGGGGAAACCTTTGGG	840
841	ATCAGGTGAAGGCCATCCATGTGAACCATGATTCGGGCTCACCTGGAGCTGGAAGAC	900
841	ATCAGGTGAAGGCCATCCATGTGAACCATGATTCGGGCTCACCTGGAGCTGGAAGAC	900
901	AAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAGTACCTACTCCA CAAACAGGACCATG	960
901	AAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAGTACCTACTCCA CAAACAGGACCATG	960
961	ATTCGGATTCCTTTGGCCTTTGTGTTCCGTGGGAGGAAAGACACCGGATATTACACC	1020
961	ATTCGGATTCCTTTGGCCTTTGTGTTCCGTGGGAGGAAAGACACCGGATATTACACC	1020
1021	TGCTCTTCTCAAAGCACCCAGCCAGTACAGCTTTGGTGACCATCTCTAGAAAAGGGTTT	1080
1021	TGCTCTTCTCAAAGCACCCAGCCAGTACAGCTTTGGTGACCATCTCTAGAAAAGGGTTT	1080
1081	ATAAACGCTACCACTCGCAAGAAAGATGAAATTGACCCCGTACGAAAAGTTCTGCTTC	1140
1081	ATAAACGCTACCACTCGCAAGAAAGATGAAATTGACCCCGTACGAAAAGTTCTGCTTC	1140
1141	TCAGTCAGGTTTAAAGGTCACCAAGATCCGATGCGAGTGGATCTTCTCTCAAGCCTCA	1200
1141	TCAGTCAGGTTTAAAGGTCACCAAGATCCGATGCGAGTGGATCTTCTCTCAAGCCTCA	1200
1201	TTTCTTGTGAACAGAGGCGCTGGAGGTGGGTACGACATATCTAAATTTGGATCAT	1260
1201	TTTCTTGTGAACAGAGGCGCTGGAGGTGGGTACGACATATCTAAATTTGGATCAT	1260
1261	AAGAAACAGGCGAGGAGTACATATTTCTATGCAAAAATGATGACCGCCAGTTCACCAA	1320
1261	AAGAAACAGGCGAGGAGTACATATTTCTATGCAAAAATGATGACCGCCAGTTCACCAA	1320



Query Match 64.5%; Score 2227.2; DB 9; Length 3501;									
Best Local Similarity 79.6%; Pred. Mismatches 0; No. Matches 2775; Conservative 0; Indels 56; Gaps 10;									
QY	8	GGCTACCGCGCTCCGGAGCCATCGGGCGTTGGCGCAGCGACGACCGGGCGGTGC	67						
DB	35	GGGACCCCGGGCTCCGGAGCCATGCGGCGTTGGC---GCGCGACGCGGGCACCGGTGC	91						
QY	68	TGCTGCTTGTGTTTGTTCAGTAATGATCTTTGAGACGGTTTACAAAACAAGACCTGCGTG	127						
DB	92	CGCTGCTGTTGTTTCTGCAATGATATTTGGGACTATTACAAATCAAGATCTGCGTG	151						
QY	128	TGATCAAGTGTGTTTAAATCAGTCATGAGAAACAATGGCTCATCAGCGGGAAGCCATCAT	187						
DB	152	TGATCAAGTGTGTTTAAATCAATCATAAAGAACATGATTCATCAGTGGGNAAGTCATCAT	211						
QY	188	CGTACCGAATGGTCGGAGGATCCCCGAAGACTCTCAAGTGTACCCCGAGCGCCAGAGTG	247						
DB	212	CATATCCCATGGTATCAGAAATCCCGGAAGACCTCGGGTGTGCGTTGAGACCCCGAGAGT	271						
QY	248	AAGGGAACGCTATATGAAGCGGCCACCTGAGAGTGGCCGAGTCTGGGTTCATCAACCTGCG	307						
DB	272	CAGGGAAGTGTACGAAGCTGCCGCTGTGGAAGTGGATGTATCTGCTTCATCACTGCG	331						
QY	308	AAGTGCAGCTCGCCACCCGAGGGACCTTTCCTGCTCTGGGTCTTTAAGCACAGCTCCC	367						
DB	332	AAGTGTGCTCGATGCCGCCAGGGAACATTTCTCTGCTCTGGGTCTTTAAGCACAGCTCCC	391						
QY	368	TGGCTCGCAGCGCATCTTTGATTTTACAAAACAGAGGAATCGTTTCAATGGCCATCTTGA	427						
DB	392	TGAATTCGCCAGCCACATTTTGATTTTACAAAACAGAGGAGTGTGTTCCATGCTCATTTTGA	451						
QY	428	ACGTGACAGAGACCCAGSCAGGAGAAATACCTACTCCATATTTCAAGAGCGAAGCGGCCAACT	487						
DB	452	AAATGACAGAAACCCAGCTGGAGAAATACCTACTTTTTATTCAGAGTGAAGCTACCAATT	511						
QY	488	ACAAGTACTGTTACAGTGAATGTAAAGATACACAGCTGTATGTGCTTAAGGAGACCTT	547						
DB	512	ACACAATATTGTTTACAGTGAGTATAAGAAATACCCCTGCTTTACATTAAGAAAGACCTT	571						
QY	548	ACTTTAGGAAGATGGAACACAGGATGCATGCTCTGCAATCTCCGAGGGTGTTCGGAGC	607						
DB	572	ACTTTAGAAAAATGGAAAAACGAGAGCCCTGCTGTGCATATCTGAGAGCGTTCAGAGC	631						
QY	608	CCACTGTGAGTGGGTGCTCTGAGCTCCACAGGGGAAGAGCTGTAAAGAAAGGCGCCTG	667						
DB	632	CGATGTGGAATGGGTGCTTTTGGCATTCACAGGGGGAAGCTGTAAAGAAAGAAAGTCCAG	691						
QY	668	CTGTGTGCAGAAAGGAGGAAAAGTACTTCATGAGTTGTTCCGAAACAGACATCAGAGTCT	727						
DB	692	CTGTGTGTTAAAAAGGAGGAAAAAAGTCTTCATGAAATTTATTTGGGACGGACATAGGTGCT	751						
QY	728	GTGCTGAAGATGCATGGGCGCGAATGCAACCAAGCTGTTCACCATAGATCTAAAAACAGG	787						
DB	752	GTGCCAGAAATGAATCGGCGAGGGAATGCACAGGCTGTTCACAATAGATCTAAATCAAA	811						
QY	788	CTCCTCAGAGGACACTGCCCCAGTATTCCTGAAAGTGGGGAAACCCTTGTGGATCAGGT	847						
DB	812	CTCCTCAGACCAATTCGCCACAAATTTATTTCTTAAAGTAGGGGAAACCCTTATGGATAAGGT	871						
QY	848	GTAAGGCCATCCATGTGAACCATGGATTCGGGCTCACCTGGGAGCTGGAAGACAAAGCCC	907						
DB	872	GCAAGCTGTTATGTGAACCATGGATTCGGGCTCACCTGGGAAATTAGAAAAACAAAGCAC	931						
QY	908	TGAGGAGGGCAGCTACTTTGATGATGAGTACCTACTCTCCAAACAGGAACCATGATTCGGA	967						
DB	932	TGAGGAGGGCAGCTACTTTGATGATGAGTACCTACTCTCCAAACAGGAACCTATGATACGGA	991						
QY	968	TTCTCTTGGCCTTTGTGCTCTTCGTGGGAAGGAACGACACCGGATATTACACTGCTCTTT	1027						
DB	992	TTCTGTTTCTTTTGTATCATCAGTGGCAGAAACGACACCGGATCTACACTGTTGCTCT	1051						





Db 3206 CTTTGTAAATCAATCATCTGTACAGGAGGAGGAGTGAATGAATGTTATTG 3265  
Qy 3210 TGAGCCCGCTACCTGGGGCTTTCCACAGCTTCAGGGGAAGCCATGATCTGAAA 3269  
Db 3266 GAGCATGATCTGCATCAAGGCTTCTCAGGCGGCTTGAGTGAATGTTGTTACCTGAAG 3325  
Qy 3270 TATAGTATATCTTTGTAATACGTGAAAACAAACCCGTTTGTGCTAAGGGAAGC 3329  
Db 3326 TACAGTATATCTTTGTAATACATAAAACAA-----AGCATTTGCTAAGGAGAGC 3378  
Qy 3330 TAAATATGATTTTAAATATCTATGTTTAAATATCTATGTAATCTTTTCAATCTATTAG 3389  
Db 3379 TAATATGATTTT--AGTCTATGTTTAAATATCTATGTAATCTTTTCAAGTATTAG 3436  
Qy 3390 TGATATATTTTATGGATGAATAAATCTTCTACTGTAAATAAATAAATAAATAAATAA 3449  
Db 3437 TGATATATTTTATGGTGGGAATAAATTTCTACTACAGTAAATAAATAAATAAATAA 3496  
Qy 3450 AAAA 3453  
Db 3497 AAAA 3500

## RESULT 4

US-09-872-136-3  
; Sequence 3, Application US/09872136  
; Patent No. US20020119545A1

## GENERAL INFORMATION:

APPLICANT: Lemischka, Ihor R.  
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
RECEPTORS AND THEIR LIGANDS

## NUMBER OF SEQUENCES: 10

## CORRESPONDENCE ADDRESS:

ADDRESSEE: Imclone Systems Incorporated  
STREET: 180 Varick Street  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10014

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/872,136  
FILING DATE: 01-Jun-2001  
CLASSIFICATION: <Unknown>

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/208,786  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US/09/021,324  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US/07/977,451  
FILING DATE: 1992-11-19  
APPLICATION NUMBER: US/07/906,397  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: US PCT/US92/05401  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: TW 81102961  
FILING DATE: 15-APR-1992  
APPLICATION NUMBER: US PCT/US92/02750  
FILING DATE: 02-APR-1992  
APPLICATION NUMBER: US/07/813,593  
FILING DATE: 24-DEC-1991  
APPLICATION NUMBER: US/07/793,065  
FILING DATE: 15-NOV-1991  
APPLICATION NUMBER: US/07/728,913  
FILING DATE: 28-JUN-1991  
APPLICATION NUMBER: US/07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Felt, Irving N.

REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054

## INFORMATION FOR SEQ ID NO: 3:

## SEQUENCE CHARACTERISTICS:

LENGTH: 3501 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

## FEATURE:

NAME/KEY: CDS

LOCATION: 58..3039

## FEATURE:

NAME/KEY: mat\_peptide

LOCATION: 139..3036

## FEATURE:

NAME/KEY: sig\_peptide

LOCATION: 58..138

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-872-136-3

## Query Match

Best Local Similarity 64.5%; Score 2227.2; DB 10; Length 3501;

Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;

Qy	8	GGCTACCGCGGCTCCGAGGCGCATGGCGGGGTTGGCGCAGCGCAGCGCGCGCTGC	67
Db	35	GGGACCGCGGCTCCGAGGCGCATGGCGGGGTTGGCGCAGCGCAGCGCGCGCTGC	91
Qy	68	TGCTGCTTGTGTTTGTGTAATGATTTCTTGACACCGTTTACAAACCAAGACCTGCCTG	127
Db	92	CGCTGCTGTTGTTTCTGCAATGATATTTGGGACTATTACAAATCAAGATCTGCCTG	151
Qy	128	TGATCAAGTGTGTTTAAATCAGTCATGAGAACATGGCTCATCAGCGGGAAGCCATCAT	187
Db	152	TGATCAAGTGTGTTTAAATCAGTCATGAGAACATGGCTCATCAGCGGGAAGCCATCAT	211
Qy	188	CGTACCGAATGTGCGAGGATCCCGAGAGACCTCCAGTGACCCCGAGGCGCCAGAGTG	247
Db	212	CATATCCATGGTATCAGAACTCCCGAGAGACCTCGGGTGGCTTGAGACCCAGAGCT	271
Qy	248	AAGGACGGTATATGAAAGCGGCCACCGTGAGGTGGCCGAGTCTGGGTCCATCACCTGC	307
Db	272	CAGGACAGTGTACGAAGCTCCCGCTGTGGAAGTGGATGTATCTGCTTCCATCACCTGC	331
Qy	308	AAGTGACGCTCGCCACCGCGGAGACCTTTCCTGCTCTGGGCTCTTAAAGCAGCTCCC	367
Db	332	AAGTGCTGGTGGATGTCGCGGAGGAACTTTCCTGCTCTGGGCTCTTAAAGCAGCTCCC	391
Qy	368	TGGGCTGCCAGCGCACTTTGATTTACAAAACAGAGGAATCGTTTCCATGCGCATCTTGA	427
Db	392	TGAATTGCCAGCCACATTTGATTTACAAAACAGAGGATGTTTCCATGCGCATCTTGA	451
Qy	428	ACGTGACAGAGACCCAGGAGGAGAAATACCTACTTCCATATTTACAGAGCGAAACCGCCAACT	487
Db	452	AAATGACAGAAACCCAGAGTGGAGAAATACCTACTTTTATTCAGAGTGAAGTACCAATT	511
Qy	488	ACACAGTACTGTTTACAGTGAATGTAAGAGATACACAGCTGTATGTCTAAGAGACCTT	547
Db	512	ACACAATATTTTACAGTGAATGTAAGAGATACACAGCTGTATGTCTAAGAGACCTT	571
Qy	548	ACTTTAGGAAGATGGAACACCCAGGATGCACTGCTGTCATCTCCGAGGTGTTCCGAGC	607
Db	572	ACTTTAGGAAGATGGAACACCCAGGATGCACTGCTGTCATCTCCGAGGTGTTCCGAGC	631
Qy	608	CCACTGTGAGTGGGTGCTCTGCACTCCCAAGGGAAGCTGTAAGAGAGAGGCCCTG	667







```
QY 2868 AGGATGTCAGTGGCAGAGGAGAGAG-----CATGTATC 2904
Db |||||||
QY 2829 AGGATGTCAGTGGCAGATGAGAGAGAGCGATGATCAGAAATGTGGATGCCGTGTTTC 2888
Db |||||||
QY 2905 AGAATCCATCCATCTACCAAAAACAGCGCGCCCTCAGCAGAGAGAGCGGCGCTCAGAGC- 2963
Db |||||
QY 2889 GGAATGTCCTCACACCTTACCAAAACAGCGGACCTTTTACAGCAGAGAGATGGATTTGGGGCT 2948
QY 2964 CAGTCCGACAGCGCCAGGT-GAAGATTTCACAGAGAAAGATTAGCGAGAGGCGCTTG 3022
Db ACTCTCTCCGAGGCTCAGTTCGAAGATTTCGTAGAGGAAATTTAGTTTAAAGGACTTC 3008
QY 3023 GACCC-----CGCCACCTTAGCAGGCTGTAGACCGCAGCCAAAGATTAGCCCTCCCTCT- 3077
Db |||||
QY 3009 ATCCCTCCACTATCCCTTAACAGGCTGTAGATTACCAAAAACAAAGGTTAAATTTTCATCACTA 3068
QY 3078 -GAGGAAGCGCCCTACACGCGGTTGCTTGGCTGGACTTTTCTCTAGA 3123
Db |||||||
QY 3069 AAAGAAATCTATATCACTGCTGCTTACCAGACTTTTCTCTAGA 3115
Db |||||||

RESULT 6
US-09-796-692-7538
; Sequence 7538, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 2077-001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7538
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-7538

Query Match 13.5%; Score 466.6; DB 10; Length 645;
Best Local Similarity 84.5%; Pred. No. 5.1e-115;
Matches 550; Conservative 0; Mismatches 94; Indels 7; Gaps 2;

QY 2035 CTGGGACACCATGACACATCTGATCTGCTGGGGGCGATGACATGTCAGGGCAGTG 2094
|||
```

```
Db 1 CTGGGAAGCCAGAGAAATATTGTGAACCTGCTGGGGCGGTGCACACTGTCTAGGACCAAT 60
QY 2095 TACTTGATTTTGAATATTGCTATGTGTGACCTCTCAACTACCTAAAGAAAGTAAAGA 2154
|||
Db 61 TACTTGATTTTGAATATTGCTATGTGTGATCTTCTCAACTATCTAAGAAAGTAAAGA 120
|||
QY 2155 GAGAGTTTTCACAGGACATGACAGAGATTTTAAAGAACATAATTTTCAGTTCTTACCC 2214
|||
Db 121 GAAAAATTTTCACAGGACTTGGACAGAGATTTTCAAGGAAACACAATTTTCAGTTTACCCC 180
|||
QY 2215 ACTTTCCAGGACCAATTCAAATTCAGCATGCTGTGTTTACAGAGAGTTTCAGTTACCCG 2274
|||
Db 181 ACTTTCCATCAATCAATTCAAATTCAGCATGCTGTGTTTCAAGAGAGTTTCAGATACCCC 240
|||
QY 2275 CCCTTGATCAGCTCTCAGGGTTCAATGGGAATTCAAATTCATCTGAAGATGAGATTGAA 2334
|||
Db 241 GACTCGGATCAATCTCAGGGCTTCATGGGAATTCATTTCACTGAGATGAAATTTGAA 300
|||
QY 2335 TATGAAAAACAGAGAGGCTGGCAGAGAAAGAGAGAGGAGAAATTTGAACTGTCACGTTT 2394
|||
Db 301 TATGAAAAACCAAAAAAGGCTG-----GAAGAGAGAGGAGACTTGAATGTGCTTACATTT 354
|||
QY 2395 GAAGACCTCTTGTGTTTGGTACCAAGTGGCCAAAGGCATGGAAATTCCTGGAGTTCAAG 2454
|||
Db 355 GAAGATCTTCTTTGCTTTGCAATATCAAGTTGCCAAAGGAATGGAATTTCTGGAAATTAAG 414
|||
QY 2455 TCCTGTGTCCACAGAGACCTGGCAGCAGGAATGTGTGTGTCCACCCAGGGAAGTGTG 2514
|||
Db 415 TCCTGTGTCCACAGAGACCTGGCGCAGGAACGTGCTGTCCACCCAGGGAAGTGTG 474
|||
QY 2515 AGATCTGTGACTTTTGGACTGGCCCGAGACATCTGAGCGACTCCAGCTACGTGTCAGG 2574
|||
Db 475 AGATATGTGACTTTGGATTTGCTCGAGATATCATGAGTGATTTCCAACTATGTTGTCAGG 534
|||
QY 2575 GGCACGCGCGCTGGCGGTGAGTGGAGTGGACCGACCCGAGAG-CCTATTGAAGGGATCTA 2633
|||
Db 535 GGCATGCCCCGCTGCTGCTGTAAATGGATGGCGCCCCGAAAGCCCTGTTTGAAGGCATCTA 594
|||
QY 2634 CACAATCAAGAGTGACGTCTGTGTCCTACGGCATCTTCTCTGGGAGATATT 2684
|||
Db 595 CACCATTAAGATGATGCTGTGTCATATGGAATATTACTGTGGGAATCTT 645
|||

RESULT 7
US-10-040-862-7538
; Sequence 7538, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
```



```

; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7538
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-7538

Query Match      13.5%; Score 466.6; DB 14; Length 645;
Best Local Similarity 84.5%; Pred. No. 5.1e-115; Mismatches 94; Indels 7; Gaps 2;
Matches 550; Conservative 0;

Qy 2035 CTGGGACACCATGACAAATCGTGAATCTGCTGGGGGATGACACACTGTGAGGGCCAGTG 2094
Db 1 CTGGGAAGCCACGAGAAATATTGTGAACCTGCTGGGGGGTGCACACTGTGAGGACCAATT 60

Qy 2095 TACTTGATTTTGAATATTGTGCTATGGTGACCTCTCACTCACTAAGAGTAAAGA 2154
Db 61 TACTTGATTTTGAATATTGTGCTATGGTGACCTCTCACTCACTAAGAGTAAAGA 120

Qy 2155 GAGAGTTTTCAGACACATGACAGAGATTTTAAAGGACATAATTTACGTTTACCCCT 2214
Db 121 GAAAAATTTTCAGACACTGTGACAGAGATTTTCAAGGAAACACAAATTTACCCCT 180

Qy 2215 ACTTTCCAGGACATTCAAATTTCCAGCATGCTGTTTCAAGAGAGTTTCAGTTACACCG 2274
Db 181 ACTTTCCAAATCACATCCAAATTTCCAGCATGCTGTTTCAAGAGAGTTTCAGTTACACCG 240

Qy 2275 CCCTTGGATCAGCTCTCAGGGTTCAATGGGAATTCATTTCACTCTGAAGTGAATTTAA 2334
Db 241 GACTCGGATCAAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAGTGAATTTAA 300

Qy 2335 TATGAAACCCAGAGAGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2394
Db 301 TATGAAACCCAGAGAGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 354

Qy 2395 GAAGACCTCTCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 2454
Db 355 GAAGATCTTCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 414

Qy 2455 TCGTGTGTCCAGAGACCTGCGACCCAGGAATGTTGGTCAACCCAGGGAGGTGGTG 2514
Db 415 TCGTGTGTTCAGAGACCTGCGCCCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 474

Qy 2515 AAGATCTGTGATTTGGATCTGCGCCGAGACATCTCTGAGGAGATCTCAGCTACGTCAGG 2574
Db 475 AAGATATGTGATTTGGATTTGGCTGCGAGATATCATGAGTATTCAGTATCTGTCAGG 534

Qy 2575 GGCACGACCGCTGCGGTTGAAGTGGATGGACCCGAGAG-CTTATTGAGGAGATCTA 2633
Db 535 GGCATGCGCGTCTGCGCTGTAATGAGTGGCCCGGAGAGGCTGTTTGAAGGAGATCTA 594

Qy 2634 CACAAATCAAGAGTACGCTCTGCTCTCAGGCACTCTCTCTGAGAGATATT 2694
Db 595 CACCAATTAAAGAGTATGCTGTGCTATATGGAATATCTGTGGGAAATCTT 645

```

```

RESULT 8
US-09-796-692-8085
; Sequence 8085, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPI
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8085
; LENGTH: 610
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-8085

Query Match      13.2%; Score 456.4; DB 10; Length 610;
Best Local Similarity 85.1%; Pred. No. 2.8e-112; Mismatches 86; Indels 6; Gaps 1;
Matches 524; Conservative 0;

Qy 2035 CTGGGACACCATGACAAATCGTGAATCTGCTGGGGGATGACACACTGTGAGGCCAGTG 2094
Db 1 CTGGGAAGCCACGAGAAATATTGTGAACCTGCTGGGGGGTGCACACTGTGAGGCCAAATT 60

Qy 2095 TACTTGATTTTGAATATTGTGCTATGGTGACCTCTCACTCACTAAGAGTAAAGA 2154
Db 61 TACTTGATTTTGAATATTGTGCTATGGTGACCTCTCACTCACTAAGAGTAAAGA 120

Qy 2155 GAGAGTTTTCAGGACATGACAGAGATTTTAAAGGAACAATAATTCAGTTCTTACCCCT 2214
Db 121 GAAAAATTTTCAGGACTGTGACAGAGATTTTCAAGGAAACAATAATTCAGTTCTTACCCCT 180

Qy 2215 ACTTTCCAGGACATTCAAATTTCCAGCATGCTGTTTCAAGAGAGTTTCAGTTACACCG 2274
Db 181 ACTTTCCAAATCACATCCAAATTTCCAGCATGCTGTTTCAAGAGAGTTTCAGTTACACCG 240

Qy 2275 CCCTTGGATCAGCTCTCAGGGTTCAATGGGAATTCATTTCACTCTGAAGTGAATTTAA 2334
Db 241 GACTCGGATCAAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAGTGAATTTAA 300

Qy 2335 TATGAAACCCAGAGAGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2394
Db 301 TATGAAACCCAGAGAGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 354

Qy 2395 GAAGACCTCTCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 2454
Db 355 GAAGATCTTCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 414

Qy 2455 TCGTGTGTCCAGAGACCTGCGACCCAGGAATGTTGGTCAACCCAGGGAGGTGGTG 2514

```

Db 415 TCCTGTGTTTACAGAGACCTGGCCGCGCAGGACGTCCTGTACCCACCGGAAAGTGGTG 474  
QY 2515 AGATCTGTGACTTTGGACTGGCCGAGACATCTGAGCGACTCCAGCTACGTCGTCAGG 2574  
Db 475 AGATATGTGACTTTGGATGGCTCGATATATCATGAGTGATTTCCAACTATGTGTGAGG 534  
QY 2575 GGCACCGCAGCGGTGCGGTGGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGGATCTAC 2634  
Db 535 GGCATGCGCGTCTGCTGTAAATGATGGCCCGCCGAAAGCCTGTTTGAAGGCATCTAC 594  
QY 2635 ACAATCAAGAGTGACG 2650  
Db 595 ACCATTAAAGAGTGATG 610

## RESULT 9

US-10-040-862-8085  
; Sequence 8085, Application US/10040862  
; Publication No. US20030078396A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; FILE REFERENCE: 014058-013520US  
; CURRENT APPLICATION NUMBER: US/10/040,862  
; CURRENT FILING DATE: 2001-11-06  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: US 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: US 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: US 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: US 60/223,416  
; PRIOR FILING DATE: 2000-08-04  
; PRIOR APPLICATION NUMBER: US 60/223,378  
; PRIOR FILING DATE: 2000-08-07  
; PRIOR APPLICATION NUMBER: US 09/796,692  
; PRIOR FILING DATE: 2001-03-01  
; NUMBER OF SEQ ID NOS: 10467  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8085  
; LENGTH: 610  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-040-862-8085  
Query Match 13.2%; Score 456.4; DB 14; Length 610;  
Best Local Similarity 85.1%; Pred. No. 2.8e-112;  
Matches 524; Conservative 0; Mismatches 86; Indels 6; Gaps 1;  
QY 2035 CTGGGACCATGACATCGTGAATCTGCTGGGGGCGATGCACACTGTGAGGGCCAGTG 2094  
Db 1 CTGGGAGCCACGAGATATTTGAACTGCTGGGGGGGTGCACACTGTGAGGCCATCTAC 594

QY 2095 TACTTGATTTTGAATATTTGCTATGCTGATGCTCTCACTACCTAAGAGTAAAGA 2154  
Db 61 TACTTGATTTTGAATATTTGCTATGCTGATGCTCTCACTATCTAAGAGTAAAGA 120  
QY 2155 GAGAAGTTTACAGGACATGGACAGAGATTTTAAAGGAACATATTTTCACTTCTTACCT 2214  
Db 121 GAAAATTTTACAGGACTTGGACAGAGATTTTCAAGGAACACATTTTCACTTTTACCCC 180  
QY 2215 ACTTTCAGGACATTTCAAATTTCCAGCATGCTGCTGTTTCAAGAGAGTTTCACTTACACCCG 2274  
Db 181 ACTTTCAGATCAATCAATTTCCAGCATGCTGCTGTTTCAAGAGAGTTTCACTTACACCCG 240  
QY 2275 CCTTGGATCAGCTCTCAGGTTTCAATGGGAATTTCAATTTCTGAAGATGAGATTGAA 2334  
Db 241 GACTCGGATCAATCTCAGGGCTTTCATGGGAATTTTCACTTCTGAAGATGAAATGAA 300  
QY 2335 TATGAAACGACAGAGAGGCTGCGAGAGAGAGGAGGAGATTTTGAACGCTGACGTTT 2394  
Db 301 TATGAAACGACAGAGAGGCTGCGAGAGAGGAGGAGATTTTGAACGCTGACGTTT 354  
QY 2395 GAAGACCTCTCTTGTCTTGTGCTTACCAAGTGGCCAAAGGATGGAATTTCTGGAGTTCAAG 2454  
Db 355 GAAGATCTCTTGTCTTGTGCTTACCAATCAAGTGGCCAAAGGATGGAATTTCTGGAGTTTAAAG 414  
QY 2455 TCGTGTGTTCCAGAGACCTGCGAGAGAGGATGTTGTTGTTCAACCGGAAAGTGGTG 2514  
Db 415 TCGTGTGTTCCAGAGACCTGCGAGAGAGGATGTTGTTGTTCAACCGGAAAGTGGTG 474  
QY 2515 AGATCTGTGACTTTGGACTGGCCGAGACATCTTGGAGGATGTTTGAAGGGATCTAC 2634  
Db 475 AAGATATGTGACTTTGGATGTTGCTCGATATATCATGAGTGATTTCCAACTATGTGTGAGG 534  
QY 2575 GGCACCGCAGCGCTGCGGTGGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGGATCTAC 2634  
Db 535 GGCATGCGCGTCTGCTGTAAATGATGGCCCGCCGAAAGCCTGTTTGAAGGCATCTAC 594  
QY 2635 ACAATCAAGAGTGACG 2650  
Db 595 ACCATTAAAGAGTGATG 610

## RESULT 10

US-09-796-692-8927  
; Sequence 8927, Application US/09796692  
; Publication No. US20020198362A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
; FILE REFERENCE: 2077.001200  
; CURRENT APPLICATION NUMBER: US/09/796,692  
; CURRENT FILING DATE: 2001-03-01  
; PRIOR APPLICATION NUMBER: 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/222,903



```
QY 2515 AAGATCTGTGACTTTGGAGTGGCCCGAGACATCTCGAGCGACTCCAGCTACGTGCTCAGG 2574
|||||
Db 475 AAGATATGTGACTTTGGATTGGCTCGAGATATCATGAGTGAATTCACCAACTATGTTGTGAGG 534
|||||
QY 2575 GCGAAGCGACGGCTGCGGGTGAAGTGAATGGACCGAGAGCTTATTGTAAGGATCTAC 2634
|||||
Db 535 GCGAATGCCGCTGCTGCTGTAATGATGGCCCGAGAGCTGTTTGAAGGATCTAC 594
|||||
QY 2635 ACAATCAAGAG 2645
|||||
Db 595 ACCATTAGAG 605
|||||

RESULT 12
US-09-796-692-7976
; Sequence 7976, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077, 001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7976
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-7976

Query Match 13.0%; Score 448.8; DB 10; Length 597;
Best Local Similarity 85.4%; Pred. No. 3.1e-110;
Matches 514; Conservative 0; Mismatches 82; Indels 6; Gaps 1;

QY 2035 CTGGACACCATGACAAATCTGTGCTGGGGGCGATGACACTGTGAGGGCCAGTG 2094
|||||
Db 1 CTGGAGAGCCAGAGATATTGTGAACCTGCTGGGGGGGTGACACTGTGAGGACCAATT 60
|||||

QY 2095 TACTTGATTTTCAATATTGTTGCTATGTGACCTCTCAACTTACCTAAAGAGTAAAGA 2154
|||||
Db 61 TACTTGATTTTGAATGCTGTTGCTATGTGATCTTCTCAACTATCTAAGAGTAAAGA 120
|||||

QY 2155 GAGAAGTTTCAGAGGACATGGACAGAGATTTTAAAGAACATAATTTTCAGTTTACCCT 2214
|||||
```

```
Db 121 GAAAAATTTTCAGAGGACTTGGACAGAGATTTTCAAGAAACACAATTTTCAGTTTATCCCC 180
QY 2215 ACTTTCCAGGCAATTCAAAATCCAGCATGCTCGTGTTCACGAGAAGTTTCAGTTACACCCG 2274
|||||
Db 181 ACTTTCCAATCACAATCCAAAATCCAGCATGCTCGTGTTCAGAGAAGTTTCAGATACACCCG 240
|||||
QY 2275 CCCTTGGATCAGCTCTCAGGGTTCAATGGGAATTCATTTCAATTCCTGAAGATGAGATTGAA 2334
|||||
Db 241 GACTCGGATCAAAATCTCRGGGCTTCATGGGAATTCATTTCACTCTGAAGATGAAATTGAA 300
|||||
QY 2335 TATGAAAACGAGAAGAGGCTGGCAGAGAAGAGGAGGAGGAAATTTTGAACGTGCTGACGTTT 2394
|||||
Db 301 TATGAAAACGAAAAGAGGCTG-----GAAGAAGAGAGGACTTGAATGCTTTCATATTT 354
|||||
QY 2395 GAAGACCTCTTTGCTTTGCTTACCAAGTGGCCAAAGGCAATGGAATTCCTGGAGTTCAAG 2454
|||||
Db 355 GAAGATCTTTTGTCTTTGCTTATCAAGTTGCTCAAGGAATGGAATTTCTTGGAAATTTAAG 414
|||||
QY 2455 TCGTGTGTCACAGAGACCTTGGCAGCAGGAATGTTGTTGCTACCCACGGAAGGTGGTG 2514
|||||
Db 415 TCGTGTGTTTCACAGAGACCTTGGCCCGCCAGGAACGTGTTGTCTACCCACGGAAGGTGGTG 474
|||||
QY 2515 AAGATCTGTGACTTTTGACTTGGCTGGCCCGCAGACATCTCTGAGCGACTCCAGTACGTCGTCAGG 2574
|||||
Db 475 AAGATATGTGACTTTGGATTGGCTCGAGATATCATGATGATTCACAACTATGTTGTCAGG 534
|||||
QY 2575 GGCAACGCA CGGCTGCCGCTGGAAGTGGATGGCACCAGAGCTTATTGGAAGGATCTAC 2634
|||||
Db 535 GGCAATGCCGCTGCTGCTGTAATGATGGCCCGCCGAAAGCCCTGTTTGAAGGCACTAC 594
|||||
QY 2635 AC 2636
|||
Db 595 AC 596
|||

RESULT 13
US-10-040-862-7976
; Sequence 7976, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-01352005
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
```

```

; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7976
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-7976

Query Match      13.0%; Score 448.8; DB 14; Length 597;
Best Local Similarity 85.4%; Pred. No. 3.1e-110;
Matches 514; Conservative 0; Mismatches 82; Indels 6; Gaps 1;

Qy 2035 CTGGGACACCATGACCAATCGTGAATCTGTGGGGGCATGCACACTGTCAAGGCGCAGTG 2094
Db      |||||
1 CTGGGAAGCCACGAGATAATTGTGAACCTGTCTGGGGCGTGCACTGTCAAGACCAATT 60

Qy 2095 TACTTGAATTTTGAATATTTGTCTATGTGTGACCTCTCAACTACCTAAGAAGTAAAGA 2154
Db      |||||
61 TACTTGAATTTTGAATATCTGTGTCTATGTGTCTCTCAACTATCTAAGAAGTAAAGA 120

Qy 2155 GAGAAGTTTTCACAGGACATGCACAGAGATTTTAAAGAACATATTTTCAGTTCTTACCT 2214
Db      |||||
121 GAAAAATTTTCACAGGACTTGGACAGAGATTTTCAAGGAACAAATTCAGTTTTTACCCC 180

Qy 2215 ACTTTCAGGCACATTCAAATTCACAGCATGCTCGTTTACGAGAAGTTTCAGTTACACCCG 2274
Db      |||||
181 ACTTTCCAATCACATCCAAATTCAGCATGCTCGTTTCAAGAGAAGTTTCAGTACACCCG 240

Qy 2275 CCCTTGATCAGCTCTCAGGGTTCAATGGGAATTCAAATTCATCTGAAGATGAGATTGAA 2334
Db      |||||
241 GACTCGGATCAAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAGATGAAATTGAA 300

Qy 2335 TATGAAACCCAGAGAGGCTGGCAGAGAAGAGGAGGAAGATTTGAACGTGCTGACGTTT 2394
Db      |||||
301 TATGAAACCCAAAAAGGCTG-----GAAGAAGAGGAGGACTTGAATGTGCTTACATTT 354

Qy 2395 GAAGACCTCTTTCCTTTCGTACCAAGTCGGCCAAAGGCATGGAATTCCTGGAGTTCAAG 2454
Db      |||||
355 GAAGATCTTTCCTTTCGTATCAAGTTGCCAAGGAATGGAATTCCTGGAATTTAAG 414

Qy 2455 TCGTGTGTCACAGACCTGGCAGCAGGAATGTGTGTGTCAACCCACGCGGAAGGTG 2514
Db      |||||
415 TCGTGTGTTTACAGAGACCTGGCCGCGAGGAACGTGCTTGTCAACCCACGCGGAAGGTG 474

Qy 2515 AAGATCTGTGACTTTTGGACTGGCCCGGACATCTCTGAGGACCTCAGACTACGTGTCAGG 2574
Db      |||||
475 AAGATATGTGACTTTTGGATTGGCTCGAGATATCATGAGTGAATCCAACTATGTGTGTGAGG 534

Qy 2575 GGCAACCGCAGGCTGCCGCTGAAGTGGATGGCACCCGAGAGCTTATTTGGAAGGGATCTAC 2634
Db      |||||
535 GGCAATGCCGCTCGCTGTAAATGGATGGCCCCCGAAGCCCTGTTTGAAGGCATCTAC 594

Qy 2635 AC 2636
Db      ||
595 AC 596

```

RESULT 14  
US-09-796-692-8273  
; Sequence 8273; Application US/09796692  
; Publication No. US20020198362A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
; FILE REFERENCE: 2077.001200  
; CURRENT APPLICATION NUMBER: US/09/796,692  
; CURRENT FILING DATE: 2001-03-01

```

; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8273
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (586)
; OTHER INFORMATION: n=A,T,C or G
US-09-796-692-8273

```

Query Match	13.0%;	Score 448;	DB 10;	Length 601;
Best Local Similarity	84.8%;	Pred. No. 5.1e-110;		
Matches 515;	Conservative 0;	Mismatches 86;	Indels 6;	Gaps 1;
Qy	2035	CTGGGACACCATGACAACTCGTGAATCTGCTGGGGCATGCACACTGTGAGGGCCAGTG	2094	
Db	1			
Qy	2095	TACTTGATTTTTGAAATATTGTTGCTATGTTGACCTCTCTCAACTACCTTAAGAAGTAAAGA	2154	
Db	61	TACTTGATTTTGAATACCTGTTGCTATGTTGATCTTCTCAACTATCTPAAGAAGTAAAGA	120	
Qy	2155	GAGAAGTTTCAAGGACATGGAACAGAGATTTTTTAAGSAACATAATTTCAGTCTTTACCCCT	2214	
Db	121	GAAGAAATTTCAAGGACTTGGACAGAGATTTTCAAGGAACACAATTTCAGTTTTACCCC	180	
Qy	2215	ACTTTCCAGGCACATTCMAATTCACAGATGCCGTGTTACAGAGAA GTTCAGTTTACACCG	2274	
Db	181	ACTTTCCAAATCAATCCAAATTCACGATGCCGTGTTCAAGAGAAGTTCAGATACACCCG	240	
Qy	2275	CCCTTGGATCAGCTCTCAGGGTTTCAATGGGAATTCAAATTCATTTCTGAAGATGAGATTGAA	2334	
Db	241	GACTTCGGATCNAATCTCAGGGCTTCAATGGGAATTCATTTCACTCTGAAGATGAAATTGAA	300	
Qy	2335	TATGAAAAACAGAAAGGCTGGCAGAAAGAGGAGGAAGATTGAACTGCTGACGTTT	2394	
Db	301	TATGAAAAACAAAAAGGCTG-----GAAGAAGAGGAGGACTTGAATGTGCTTACATTT	354	
Qy	2395	GAGACCTCTTTGCTTTTCGTACCAAGTGGCCAAAGGCATGGAATTCCTGAGTTTCAAG	2454	
Db	355	GAGATCTTCTTTGCTTTGCTATCAAGTTCGCAAGGAATGGAATTTCTGGAATTTAAG	414	
Qy	2455	TCTGTGTCCACAGACCTGGCAGCCAGGAATGTGTTGGTCACCCACGGGAGGTGGT	2514	
Db	415	TCGTGTGTTACAGACACTGGCCGCGAGGAACGTGCTTTGTCAACCCACGGGAAGGTGGT	474	
Qy	2515	AGATCTGTGACTTTTGGACTGCCCGCAGACATCTCTGAGCGACTCAGACTTACGTCGAGG	2574	

Db 475 AAGATATGACCTTTGGATGGCTCGATATATCATGAGTGATCCAACTATGTTGTCAGG 534  
Qy 2575 GGCAACGACGCGCTCGCGTGGAAGTGGATGGCACCAGAGCTTATTTGAAGGGATCTAC 2634  
Db 535 GGCAATGCCGCTCGCTGTAAATGATGGGCGCCCGAAAGCCTGTTTGAANGCATCTAC 594  
Qy 2635 ACAATCA 2641  
Db 595 ACCATTA 601

RESULT 15  
US-10-040-862-8273  
; Sequence 8273, Application US/10040862  
; Publication No. US20030078396A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; FILE REFERENCE: 014058-013520US  
; CURRENT APPLICATION NUMBER: US/10/040,862  
; PRIOR FILING DATE: 2001-11-06  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: US 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: US 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: US 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: US 60/223,416  
; PRIOR FILING DATE: 2000-08-04  
; PRIOR APPLICATION NUMBER: US 60/223,378  
; PRIOR FILING DATE: 2000-08-07  
; PRIOR APPLICATION NUMBER: US 09/796,692  
; PRIOR FILING DATE: 2001-03-01  
; NUMBER OF SEQ ID NOS: 10467  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8273  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (586)  
; OTHER INFORMATION: n=A,T,C or G  
US-10-040-862-8273

Query Match 13.0%; Score 448; DB 14; Length 601;  
Best Local Similarity 84.8%; Pred. No. 5.1e-110;  
Matches 515; Conservative 0; Mismatches 86; Indels 6; Gaps 1;  
Qy 2035 CTGGGACACCATGACAACTCGTGAATCTGTGGGGGCGATGCACACTGTGACGGCCAGTG 2094  
Db 1 CTGGGAAGCCAGAGAAATATTGAACTCTGCTGGGGGCGTGACACTGTGACGGACCAATT 60

Qy 2035 TACTTGAATTTTGAATATTGTTGCTATGCTGACCTCCTCAACTACTAAGAAAGTAAAGA 2154  
Db 61 TACTTGAATTTTGAATATTGTTGCTATGCTGATCTCTCAACTACTAAGAAAGTAAAGA 120  
Qy 2155 GAGAAGTTTTCACAGGACA TGGACAGAGATTTTAAAGGAACAATAATTTCAAGTTCTTACCT 2214  
Db 121 GAAAAATTTTCACAGGACTTGGACAGAGATTTTCAAGGAACAATAATTTCAAGTTTACCCC 180  
Qy 2215 ACTTTCAGGACACATTTCAAAATTCAGCATGCTGTTTACGAGAAAGTTTCAAGTTACCCG 2274  
Db 181 ACTTTCCAATCACATCCAAATTCAGCATGCTGTTTCAAGGAAGTTCAGATACACCCG 240  
Qy 2275 CCCTTGGATCAGCTCTCAGGGTTCAATGGGATTCAAATTTCTTGAAGATGAGTTGAA 2334  
Db 241 GACTCGGATCAAAATTCAGGGCTTCATGGGAATTCATTTCACTCTGAAGATGAAATGAA 300  
Qy 2335 TATGAAAACCAAGAGAGGCTGGCAGAGAAGAGAGGAGGAATTTGAACTGTCTGACGTTT 2394  
Db 301 TATGAAAACCAAAAGGCTG-----GAAGAGAGGAGGACTTGAATGTGCTTACATTT 354  
Qy 2395 GAAGACCTCCTTTTGGCTTACCAAGTGGCAAAGGAGTGGAAATTCCTGGAGTTCAAG 2454  
Db 355 GAAGATCTTCTTTGCTTATCAAGTTGCAATGCAAGGAATGGAATTTCTGGAATTTAAG 414  
Qy 2455 TCGTGTGTCACAGAGACTTGGCAGCAGGATGTGTTGTCACCCACGGAAGTGGTG 2514  
Db 415 TCGTGTGTCACAGAGACTTGGCAGCAGGATGTGTTGTCACCCACGGAAGTGGTG 474  
Qy 2515 AAGATCTGTGACTTTGGACTGGCCCGAGACATCTGAGCGACTCCAGCTACGTCGTCAGG 2574  
Db 475 AAGATATGTGACTTTGGATTTGGCTCGATATATCATGAGTGATTCACACTATGTTGTCAGG 534  
Qy 2575 GGCAACGACGCTCCCGTGAAGTGGATGGCAACCCGAGAGCTTATTTGAAGGGATCTAC 2634  
Db 535 GGCAATGCCGCTCTGCTGTAATAATGGATGGGCGCCCGAAAGCCTGTTTGAANGCATCTAC 594  
Qy 2635 ACAATCA 2641  
Db 595 ACCATTA 601

Search completed: August 28, 2003, 13:26:34  
Job time : 1016.5 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 15:21:07 ; Search time 280.55 Seconds

(without alignments)

5432.528 Million cell updates/sec

Title: US-09-919-408A-1

Perfect score: 3453

Sequence: 1 GCGGCTGCTACCGCGC.....AAAAAAAAAAAAAAAAAAAA 3453

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:\*

1: /cgn2\_6/ptodata/2/ina/5A COMB.seq.\*

2: /cgn2\_6/ptodata/2/ina/5B COMB.seq.\*

3: /cgn2\_6/ptodata/2/ina/6A COMB.seq.\*

4: /cgn2\_6/ptodata/2/ina/6B COMB.seq.\*

5: /cgn2\_6/ptodata/2/ina/PCTUS COMB.seq.\*

6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3453	100.0	3453	1	US-07-813-593-1
2	3453	100.0	3453	1	US-07-977-451-1
3	3453	100.0	3453	1	US-07-946-507-1
4	3453	100.0	3453	1	US-08-252-517-1
5	3453	100.0	3453	1	US-07-906-397A-1
6	3453	100.0	3453	2	US-08-601-891-1
7	3453	100.0	3453	2	US-09-021-324-1
8	3453	100.0	3453	5	PCT-US92-02750-1
9	3453	100.0	3453	5	PCT-US92-05401-1
10	3453	100.0	3453	5	PCT-US92-09893-1
11	3344	96.8	3521	1	US-08-222-299-1
12	3344	96.8	3521	1	US-08-434-878-1
13	3344	96.8	3521	5	PCT-US95-03718-1
14	3344	96.8	3521	5	US-07-977-451-3
15	2227.2	64.5	3501	1	US-08-252-517-3
16	2227.2	64.5	3501	1	US-07-906-397A-3
17	2227.2	64.5	3501	1	US-08-601-891-3
18	2227.2	64.5	3501	2	US-09-021-324-3
19	2227.2	64.5	3501	5	PCT-US92-05401-3
20	2227.2	64.5	3501	5	PCT-US92-09893-3
21	2204.8	63.9	3476	1	US-08-183-211-1
22	2204.8	63.9	3476	5	PCT-US95-00176A-1
23	2203	63.8	3475	1	US-08-222-299-3
24	2203	63.8	3475	2	US-08-434-878-3
25	2203	63.8	3475	5	PCT-US95-03718-3
26	2081.8	60.3	3120	1	US-08-222-616-22
27	2081.8	60.3	3120	4	US-08-446-648-22

28 2081.8 60.3 3120 5 PCT-US95-04228-22 Sequence 22, Appli

29 1206 34.9 1894 3 US-07-912-122-3 Sequence 3, Appli

30 1206 34.9 1894 3 PCT-US93-06404-3 Sequence 3, Appli

31 329 9.5 5084 1 US-08-306-691B-21 Sequence 21, Appli

32 329 9.5 5084 5 PCT-US93-06251-25 Sequence 25, Appli

33 286.4 8.3 4054 1 US-08-180-195-35 Sequence 35, Appli

34 286.4 8.3 4054 1 US-08-477-329-35 Sequence 35, Appli

35 286.4 8.3 4054 2 US-08-475-458-35 Sequence 35, Appli

36 286.4 8.3 4054 3 US-08-980-400-35 Sequence 35, Appli

37 286.4 8.3 4054 3 US-09-583-459A-35 Sequence 35, Appli

38 286.4 8.3 4054 3 US-09-583-210-35 Sequence 35, Appli

39 286.4 8.3 4054 4 US-09-583-449A-35 Sequence 35, Appli

40 286.4 8.3 4054 4 US-09-435-059-35 Sequence 35, Appli

41 286.4 8.3 4100 1 US-08-168-917-3 Sequence 3, Appli

42 286.4 8.3 4100 2 US-08-460-510-3 Sequence 3, Appli

43 286.4 8.3 4100 2 US-08-460-490-3 Sequence 3, Appli

44 286.4 8.3 4100 5 PCT-US92-00730-3 Sequence 3, Appli

45 286.4 8.3 4100 5 PCT-US92-00862-3 Sequence 3, Appli

## ALIGNMENTS

RESULT 1

US-07-813-593-1

; Sequence 1, Application US/07813593

; Patent No. 5185438

; GENERAL INFORMATION:

; APPLICANT: Lemischka, Ihor R.

; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL

; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: IMCLONE SYSTEMS INCORPORATED

; STREET: 180 VARICK STREET

; CITY: NEW YORK

; STATE: NEW YORK

; COUNTRY: U.S.A.

; ZIP: 10014

; COMPUTER READABLE FORM: disk

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/813,593

; FILING DATE: 19920415

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/793,065

; FILING DATE: 15-NOV-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/728,913

; FILING DATE: 28-JUN-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/679,666

; FILING DATE: 02-APR-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Feit, Irving N.

; REGISTRATION NUMBER: 28,601

; REFERENCE/DOCKET NUMBER: LEM-3-PPP

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-645-1405

; TELEFAX: 212-645-2054

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3453 base pairs

; TYPE: NUCLEIC ACID

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: CDS



2041 QY CACCATGACAAATCGTGAATCTGCTGGGGGATGCACACTGCTCAGGCGCCAGTGTACTTG 2100  
2041 Db CACCATGACAAATCGTGAATCTGCTGGGGGATGCACACTGCTCAGGCGCCAGTGTACTTG 2100  
2101 QY ATTTTGAATATTTGTTGCTATGCTGACCTCTCAACTACCTAAAGAAAGTAAAGAGAGAAG 2160  
2101 Db ATTTTGAATATTTGTTGCTATGCTGACCTCTCAACTACCTAAAGAAAGTAAAGAGAGAAG 2160  
2161 QY TTTCACAGGACATGGACAGAGATTTTAAAGAAAGTAAAGTTCCTTCACTTCTTCACTTCTT 2220  
2161 Db TTTCACAGGACATGGACAGAGATTTTAAAGAAAGTAAAGTTCCTTCACTTCTTCACTTCTT 2220  
2221 QY CAGGACATTCCTTCACTGATGCTGCTTCAAGAGAGTTCAGTTCACCCGCGCTTG 2280  
2221 Db CAGGACATTCCTTCACTGATGCTGCTTCAAGAGAGTTCAGTTCACCCGCGCTTG 2280  
2281 QY GATCAGCTCTCAGGGTTCCTTCAAGAGAGTTCCTTCAAGAGAGTTCCTTCAAGAGAGTTC 2340  
2281 Db GATCAGCTCTCAGGGTTCCTTCAAGAGAGTTCCTTCAAGAGAGTTCCTTCAAGAGAGTTC 2340  
2341 QY AACGAGAGAGCTGGCAGAGAGAGAGAGAGAGAGTTCCTTCAAGAGAGTTCCTTCAAGAG 2400  
2341 Db AACGAGAGAGCTGGCAGAGAGAGAGAGAGAGAGTTCCTTCAAGAGAGTTCCTTCAAGAG 2400  
2401 QY CTCCTTTGCTTTCCTTCAAGAGAGTTCCTTCAAGAGAGTTCCTTCAAGAGAGTTCCTTCA 2460  
2401 Db CTCCTTTGCTTTCCTTCAAGAGAGTTCCTTCAAGAGAGTTCCTTCAAGAGAGTTCCTTCA 2460  
2461 QY GTCCACAGAGACCTGGCAGAGAGAGTTCCTTCAAGAGAGTTCCTTCAAGAGAGTTCCTTCA 2520  
2461 Db GTCCACAGAGACCTGGCAGAGAGAGTTCCTTCAAGAGAGTTCCTTCAAGAGAGTTCCTTCA 2520  
2521 QY TGTGACTTTGAGCTGGCCGCGACATCTGAGCGACTCCAGCTACGCTCAGGGGCGAAC 2580  
2521 Db TGTGACTTTGAGCTGGCCGCGACATCTGAGCGACTCCAGCTACGCTCAGGGGCGAAC 2580  
2581 QY GCACGGCTGCGCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2640  
2581 Db GCACGGCTGCGCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2640  
2641 QY AAGAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2700  
2641 Db AAGAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2700  
2701 QY CCTTACCCTGCGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760  
2701 Db CCTTACCCTGCGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760  
2761 QY ATGGAGAGCCATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2820  
2761 Db ATGGAGAGCCATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2820  
2821 QY TTTGACTCAAGAGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2880  
2821 Db TTTGACTCAAGAGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2880  
2881 QY GCAGGAG 2940  
2881 Db GCAGGAG 2940  
2941 QY CAGCAG 3000  
2941 Db CAGCAG 3000  
3001 QY AGAGTACGAG 3060  
3001 Db AGAGTACGAG 3060  
3061 QY AGATTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3120  
3061 Db AGATTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3120  
3121 QY AGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3180

3121 Db AGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3180  
3181 QY CGCGAG 3240  
3181 Db CGCGAG 3240  
3241 QY AGCTTGAAG 3300  
3241 Db AGCTTGAAG 3300  
3301 QY ACCAAACCCGTTTTTGTCTAAGGGGAAAGCTAAATATGATTTTAAATAATCTATGTTTTAA 3360  
3301 Db ACCAAACCCGTTTTTGTCTAAGGGGAAAGCTAAATATGATTTTAAATAATCTATGTTTTAA 3360  
3361 QY AATCATATGTAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3420  
3361 Db AATCATATGTAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3420  
3421 QY TACTCTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3453  
3421 Db TACTCTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3453

## RESULT 2

US-07-977-451-1  
; Sequence 1, Application US/07977451  
; Patent No. 5270458  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: ImClone Systems Incorporated  
; STREET: 180 Varick Street  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/977,451  
; FILING DATE: 19921119  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US UNASSIGNED  
; FILING DATE: 12-NOV-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/906,397  
; FILING DATE: 26-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: TW 81102961  
; FILING DATE: 15-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US PCT/US92/05401  
; FILING DATE: 26-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: TW 81102961  
; FILING DATE: 15-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US PCT/US92/02750  
; FILING DATE: 02-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/813,593  
; FILING DATE: 24-DEC-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/793,065  
; FILING DATE: 15-NOV-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/728,913  
; FILING DATE: 28-JUN-1991

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Feit, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3453 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 112..3006  
NAME/KEY: sig\_peptide  
LOCATION: 31..111  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 31..3009

US-07-977-451-1

Query Match 100.0%; Score 3453; DB 1; Length 3453;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CGGGCTGTGCTACCGCGGCTCCGAGGCGCATGCGGGCGTTGGCGCAGCGCAGCGACCGG	60
Db	1	CGGGCTGTGCTACCGGGGCTCGGAGGCGCATGCGGGCGTTGGCGCAGCGCAGCGACCGG	60
Qy	61	CGGCTGTGCTGCTGTTGTTTGTGAGTAATGATTTGAGACCGTTTACAAACCAAGAC	120
Db	61	CGGCTGTGCTGCTGTTGTTTGTGAGTAATGATTTGAGACCGTTTACAAACCAAGAC	120
Qy	121	CTGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAACATGGCTCATCAGCGGGAAG	180
Db	121	CTGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAACATGGCTCATCAGCGGGAAG	180
Qy	181	CCATCATCGTACCGAATGTGTGAGGATCCCGAGAGACCTCCAGTGTACCCCGAGGCGC	240
Db	181	CCATCATCGTACCGAATGTGTGAGGATCCCGAGAGACCTCCAGTGTACCCCGAGGCGC	240
Qy	241	CAGAGTGAAGGACGGTATATGAAGCGGCCACCGTGGAGGTGGCGAGTCTGGGTCCATC	300
Db	241	CAGAGTGAAGGACGGTATATGAAGCGGCCACCGTGGAGGTGGCGAGTCTGGGTCCATC	300
Qy	301	ACCTGCAAGTGCAGCTGCCACCCAGGGGACCTTTCTGCTCTGGTCTTTAAGAC	360
Db	301	ACCTGCAAGTGCAGCTGCCACCCAGGGGACCTTTCTGCTCTGGTCTTTAAGAC	360
Qy	361	AGCTCCCTGGGTGCGCAGCGGACCTTTGATTTACAAACAGAGGAATCGTTTCCATGGC	420
Db	361	AGCTCCCTGGGTGCGCAGCGGACCTTTGATTTACAAACAGAGGAATCGTTTCCATGGC	420
Qy	421	ATCTTGAACGTGACAGACCCAGGACGAGAAATACCTACTCATATTCAGAGCGAACGC	480
Db	421	ATCTTGAACGTGACAGACCCAGGACGAGAAATACCTACTCATATTCAGAGCGAACGC	480
Qy	481	GCCAATACAGACTGTTTACAGTAATGTAAGATACACAGCTGTATGTCTAAG	540
Db	481	GCCAACTACAGACTGTTTACAGTAATGTAAGATACACAGCTGTATGTCTAAG	540
Qy	541	AGACCTACTTTAGGAATGGAAACAGAGATGCACTGCTCTGATCTCCGAGGTTT	600
Db	541	AGACCTACTTTAGGAATGGAAACAGAGATGCACTGCTCTGATCTCCGAGGTTT	600

Db	541	AGACCTTACTTTAGGAAGATGGAAACACAGGATGCACTGCTCTGATCTCCGAGGTTT	600
Qy	601	CCGAGAGCCCACTGTGTGAGTGGGTGCTCTGAGCTCCCAAGGAAAGCTGTAAAGAGAA	660
Db	601	CCGAGAGCCCACTGTGTGAGTGGGTGCTCTGAGCTCCCAAGGAAAGCTGTAAAGAGAA	660
Qy	661	GGCCCTGTGTTGTCAAGAGGAGGAAAGGTACTTCAATGAGTTGTTGGGAAACAGACATC	720
Db	661	GGCCCTGTGTTGTCAAGAGGAGGAAAGGTACTTCAATGAGTTGTTGGGAAACAGACATC	720
Qy	721	AGATGCTGTGTAGAAATGCATCTGGGCCCGGAATGCCAAAGCTGTTCCACATAGATCTA	780
Db	721	AGATGCTGTGTAGAAATGCATCTGGGCCCGGAATGCCAAAGCTGTTCCACATAGATCTA	780
Qy	781	AACAGGCTCTCAGAGACACACTGCCCCAGTTATCTGAAAGTGGGGAAACCTTTGTGG	840
Db	781	AACAGGCTCTCAGAGACACACTGCCCCAGTTATCTGAAAGTGGGGAAACCTTTGTGG	840
Qy	841	ATCAGGTGAAGGCCCATCATGTGAACATGATTCGGGCTCACCTGGGAGCTGGAGAC	900
Db	841	ATCAGGTGAAGGCCCATCATGTGAACATGATTCGGGCTCACCTGGGAGCTGGAGAC	900
Qy	901	AAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAGTACCTCTCCACAAACAGGACCATG	960
Db	901	AAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAGTACCTCTCCACAAACAGGACCATG	960
Qy	961	ATTCGGAATCTCTTGGCTTTGTCTTCCGTTGGGAAAGACGACACCGGATATTACACC	1020
Db	961	ATTCGGAATCTCTTGGCTTTGTCTTCCGTTGGGAAAGACGACACCGGATATTACACC	1020
Qy	1021	TGCTCTTCTCAAGCACCCAGCCAGTCAGGCTGGTGGTGAACATCTAGAAAAAGGTTT	1080
Db	1021	TGCTCTTCTCAAGCACCCAGCCAGTCAGGCTGGTGGTGAACATCTAGAAAAAGGTTT	1080
Qy	1081	ATAAACGGTACCAGCTCGCAAGAGATGAAATTTGACCCGTACCAAAAGTTCTGCTTC	1140
Db	1081	ATAAACGGTACCAGCTCGCAAGAGATGAAATTTGACCCGTACCAAAAGTTCTGCTTC	1140
Qy	1141	TCAGTCAGGTTTAAAGCGTACCCAGAAATCCGATCGACGTGGATCTTCTCAAGCCTCA	1200
Db	1141	TCAGTCAGGTTTAAAGCGTACCCAGAAATCCGATCGACGTGGATCTTCTCAAGCCTCA	1200
Qy	1201	TTTCTCTGTAACAGAGAGGCTGGAGGATGGGTACAGCATATCTAAATTTTGGGATCAT	1260
Db	1201	TTTCTCTGTAACAGAGAGGCTGGAGGATGGGTACAGCATATCTAAATTTTGGGATCAT	1260
Qy	1261	AAGAACAGCCAGGAGAGTACATATTTCTATGACAGAAATGATGACGCCAGTTTCAACAA	1320
Db	1261	AAGAACAGCCAGGAGAGTACATATTTCTATGACAGAAATGATGACGCCAGTTTCAACAA	1320
Qy	1321	ATGTTTCAAGCTGAATATTAAGAAAGAAACCTCAAGTGTAGCAAAATGCCTCAGCCAGCAG	1380
Db	1321	ATGTTTCAAGCTGAATATTAAGAAAGAAACCTCAAGTGTAGCAAAATGCCTCAGCCAGCAG	1380
Qy	1381	GGTCTCTGTTCTGTGATGGTACCCGCTACCCCTTTGGAACCTGGAAAGGTTTCCGAC	1440
Db	1381	GGTCTCTGTTCTGTGATGGTACCCGCTACCCCTTTGGAACCTGGAAAGGTTTCCGAC	1440
Qy	1441	AAATCTCCAAATTCACGAGGAAATCCAGAGGAGTTTGGAAATAAAGGCTAACAGA	1500
Db	1441	AAATCTCCAAATTCACGAGGAAATCCAGAGGAGTTTGGAAATAAAGGCTAACAGA	1500
Qy	1501	AAAGTGTGGCGCTGTCAGCAGTACTCTAAATATGAGTGGGCGCGGAAAGG	1560
Db	1501	AAAGTGTGGCGCTGTCAGCAGTACTCTAAATATGAGTGGGCGCGGAAAGG	1560
Qy	1561	CTTCTGTGCAAAATGCTGTGCGTACAAATTTCTATGGGACGCTTTGCGAAACCATCTTTTA	1620
Db	1561	CTTCTGTGCAAAATGCTGTGCGTACAAATTTCTATGGGACGCTTTGCGAAACCATCTTTTA	1620
Qy	1621	AACCTACAGGCGCCCTTCCCTTTTCATCCAGAACATCTCTCTATGCGGACCATTTGG	1680
Db	1621	AACCTACAGGCGCCCTTCCCTTTTCATCCAGAACATCTCTCTATGCGGACCATTTGG	1680



```

CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/813,593
FILING DATE: 24-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/793,065
FILING DATE: 15-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/728,913
FILING DATE: 28-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-PPP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3453 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 31..3009
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 31..3006
US-07-946-507-1

```

Query Match	100.0%;	Score 3453;	DB 1;	Length 3453;
Best Local Similarity	100.0%;	Prod. No. 0;		
Matches 3453;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GC	GCCTCGGCTACCGCGCTCCGGAGGCCATGCGGGCGTTGGCGCAGCGAGCGACCGG	60
Db	1	GGGGCTTGGCTACCGCGCGCTCCGGAGGCCATGCGGGCGTTGGCGCAGCGAGCGACCGG	60	
QY	61	CGGCTGCTGCTGCTTGTGTTTGTCACTAATCATTTCTTGAGACCGTTTACAAACCAAGAC	120	
Db	61	CGGCTGCTGCTGCTTGTGTTTGTCACTAATGATTTCTTGAGACCGTTTACAAACCAAGAC	120	
QY	121	CTGCTGTGTGATCAAGTGTGTTTTTAATCAGTCATCAGAACAAATGGCTCATCAGCGGGAAG	180	
Db	121	CTGCTGTGTGATCAAGTGTGTTTTTAATCAGTCATCAGAACAAATGGCTCATCAGCGGGAAG	180	
QY	181	CCATCATCGTACCGAATGGTGCAGGATCCCCAGAACCTTCAGTGTACCCGAGGCGC	240	
Db	181	CCATCATCGTACCGAATGGTGCAGGATCCCCAGAACCTTCAGTGTACCCGAGGCGC	240	
QY	241	CAGAGTGAAGGGACGGTATATGAAGCGGCCACCGTGGAGGTGGCCGAGTCTGGGTCCATC	300	
Db	241	CAGAGTGAAGGGACGGTATATGAAGCGGCCACCGTGGAGGTGGCCGAGTCTGGGTCCATC	300	
QY	301	ACCTCGAAGTGAGCTCGCCACCCAGGGAGCCTTTCTGCCTCTGGGTCTTTAAGCAC	360	
Db	301	ACCTCGAAGTGAGCTCGCCACCCAGGGAGCCTTTCTGCCTCTGGGTCTTTAAGCAC	360	
QY	361	AGCTCCCTGGCGTCGAGCGGCACTTTGATTTACAAACAGAGAAATCGTTTCCATGGCC	420	
Db	361	AGCTCCCTGGCGTCGAGCGGCACTTTGATTTACAAACAGAGAAATCGTTTCCATGGCC	420	
QY	421	ATCTTTGAACGTGACAGAGACCCAGGAGGAGAAATACCTACTCCATATTCAGAGCGACGC	480	
Db	421	ATCTTTGAACGTGACAGAGACCCAGGAGGAGAAATACCTACTCCATATTCAGAGCGACGC	480	
QY	481	GCCAACTACACAGTACTGTTTCACAGTGAATGTGAAGAGATACACAGCTGTATGCTAAGG	540	

Db	481	GC	AACT	TACA	CAGT	ACT	TGTT	TC	ACGT	GAAT	GTAA	GAGAT	A	CAC	AGCT	GTAT	GTG	CTA	AGG	540																			
QY	541	AGA	CTT	TACT	TTAG	GAAG	ATG	GA	AAAC	CAGG	ATG	CAC	TGCT	CTG	CAT	CTC	CG	AGG	GT	600																			
Db	541	AG	AC	TT	ACT	TT	AG	GA	AG	ATG	GA	AAAC	CAGG	ATG	CAC	TGCT	CTG	CAT	CTC	CG	AGG	GT	600																
QY	601	CCG	AG	CC	CA	CT	GT	GG	AG	TGG	TG	CT	CT	G	CA	GC	AGT	CC	CA	G	GG	AA	AG	CT	GT	TA	AA	GA	AA	660									
Db	601	CCG	AG	CC	CA	CT	GT	GG	AG	TGG	TG	CT	CT	G	CA	GC	AGT	CC	CA	G	GG	AA	AG	CT	GT	TA	AA	GA	AA	660									
QY	661	G	GC	CT	CT	GT	TT	GC	AG	AA	AG	GA	AA	AG	GT	AT	CT	T	GA	T	GA	T	GT	TT	CG	GA	CA	CA	TC	720									
Db	661	G	GC	CT	CT	GT	TT	GC	AG	AA	AG	GA	AA	AG	GT	AT	CT	T	GA	T	GA	T	GT	TT	CG	GA	CA	CA	TC	720									
QY	721	AG	AT	CT	GT	CT	AG	AA	TG	CA	CT	TGG	CCG	CA	TG	CA	CC	AA	GC	TG	TT	C	CA	CA	AG	CT	GT	T	C	CA	CA	TC	780						
Db	721	AG	AT	CT	GT	CT	AG	AA	TG	CA	CT	TGG	CCG	CA	TG	CA	CC	AA	GC	TG	TT	C	CA	CA	AG	CT	GT	T	C	CA	CA	TC	780						
QY	781	A	AC	AG	CT	CT	CT	AG	AG	C	AC	ACT	GC	CC	CA	GT	T	AT	T	CT	G	AA	GT	TT	CT	G	AA	GT	TT	CT	G	AA	GT	TT	CT	840			
Db	781	A	AC	AG	CT	CT	CT	AG	AG	C	AC	ACT	GC	CC	CA	GT	T	AT	T	CT	G	AA	GT	TT	CT	G	AA	GT	TT	CT	G	AA	GT	TT	CT	840			
QY	841	A	T	C	AG	T	G	T	A	G	C	C	A	T	G	T	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	900			
Db	841	A	T	C	AG	T	G	T	A	G	C	C	A	T	G	T	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	900			
QY	901	A	A	A	G	C	C	T	G	G	A	G	G	G	C	A	G	T	A	C	T	T	T	G	A	G	A	T	T	T	G	A	A	G	T	T	G	960	
Db	901	A	A	A	G	C	C	T	G	G	A	G	G	G	C	A	G	T	A	C	T	T	T	G	A	G	A	T	T	T	G	A	A	G	T	T	G	960	
QY	961	A	T	T	C	G	A	T	T	C	T	T	G	G	C	T	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	1020		
Db	961	A	T	T	C	G	A	T	T	C	T	T	G	G	C	T	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	1020		
QY	1021	T	G	C	T	T	C	T	C	A	A	G	C	A	C	C	C	A	G	C	A	G	T	G	A	C	C	A	T	C	T	C	T	A	G	A	A	A	1080
Db	1021	T	G	C	T	T	C	T	C	A	A	G	C	A	C	C	C	A	G	C	A	G	T	G	A	C	C	A	T	C	T	C	T	A	G	A	A	A	1080
QY	1081	A	T	A	A	A	C	G	T	C	A	C	G	C	T	C	G	A	A	G	A	G	A	G	A	T	G	A	A	A	T	G	A	A	A	A	A	1140	
Db	1081	A	T	A	A	A	C	G	T	C	A	C	G	C	T	C	G	A	A	G	A	G	A	G	A	T	G	A	A										





```

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/252.517
  FILING DATE: 31-OCT-1994
  CLASSIFICATION: 530
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 07/977,451
    FILING DATE: 19-NOV-1992
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 07/906,397
    FILING DATE: 26-JUN-1992
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US PCT/US92/05401
    FILING DATE: 26-JUN-1992
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: TW 81102961
    FILING DATE: 15-APR-1992
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US PCT/US92/02750
    FILING DATE: 02-APR-1992
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 07/813,593
    FILING DATE: 24-DEC-1991
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 07/728,913
    FILING DATE: 28-JUN-1991
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 07/679,666
    FILING DATE: 02-APR-1991
  ATTORNEY/AGENT INFORMATION:
    NAME: Feit, Irving N.
    REGISTRATION NUMBER: 28,601
    REFERENCE/DOCKET NUMBER: LEM-3-7P
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 212-645-1405
    TELEFAX: 212-645-2054
    INFORMATION FOR SEQ ID NO: 1:
      SEQUENCE CHARACTERISTICS:
        LENGTH: 3453 base pairs
        TYPE: nucleic acid
        STRANDEDNESS: double
        TOPOLOGY: linear
        MOLECULE TYPE: cDNA
        HYPOTHETICAL: NO
        ANTI-SENSE: NO
        FRAGMENT TYPE: N-terminal
      FEATURE:
        NAME/KEY: mat_peptide
        LOCATION: 112..3006
      FEATURE:
        NAME/KEY: sig_peptide
        LOCATION: 31..111
      FEATURE:
        NAME/KEY: CDS
        LOCATION: 31..3009
US-08-252-517-1

```

```

Query Match      100.0%; Score 3453; DB 1; Length 3453;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCTGCTACCGCGCTCGGAGGCCATCGGGCGTTGCGCAGCGCAGCGACCGG 60
Db 1 GCGGCTGCTACCGCGCTCGGAGGCCATCGGGCGTTGCGCAGCGCAGCGACCGG 60
QY 61 CGGCTGCTGCTGTTGTTTGTTCAGTAATCTTGTGAGCCGTTTCAAAACCAAGAC 120
Db 61 CGGCTGCTGCTGTTGTTTGTTCAGTAATCTTGTGAGCCGTTTCAAAACCAAGAC 120
QY 121 CTGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAACATGGCTCATCAGCGGGAAG 180

```

```

Db 121 CTGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAACATGGCTCATCAGCGGGAAG 180
QY 181 CCATCATCGTACCGGAATGGTGCAGAGATCCCGAGAAGACCTCCAGTGTACCCGAGCGC 240
Db 181 CCATCATCGTACCGGAATGGTGCAGAGATCCCGAGAAGACCTCCAGTGTACCCGAGCGC 240
QY 241 CAGAGTGAAGGGACGGGTATATGAAGCGGCCACCGTGGAGGTGGCCGAGTCTGGGTCCATC 300
Db 241 CAGAGTGAAGGGACGGGTATATGAAGCGGCCACCGTGGAGGTGGCCGAGTCTGGGTCCATC 300
QY 301 ACCCTGCAAGTGCAGCTCGCCACCCCGAGGGAACCTTCTGCTCTGGGTCTTTAAGCAC 360
Db 301 ACCCTGCAAGTGCAGCTCGCCACCCCGAGGGAACCTTCTGCTCTGGGTCTTTAAGCAC 360
QY 361 AGCTCCCTGGGTGCCAGCGGCACCTTTGATTACAAAACAGAGGAATCGTTTCATGGCC 420
Db 361 AGCTCCCTGGGTGCCAGCGGCACCTTTGATTACAAAACAGAGGAATCGTTTCATGGCC 420
QY 421 ATCTTGAACGTGACAGAGACCCAGGAGGAGATACCTACTCCATATTCAGAGCGAAGCG 480
Db 421 ATCTTGAACGTGACAGAGACCCAGGAGGAGATACCTACTCCATATTCAGAGCGAAGCG 480
QY 481 GCCAACTACACAGTACTGTTTCAAGTGAATGTAAAGATACACAGCTGTATGTCTAAGG 540
Db 481 GCCAACTACACAGTACTGTTTCAAGTGAATGTAAAGATACACAGCTGTATGTCTAAGG 540
QY 541 AGACCTTACTTTAGGAAGATGGAACACAGGATGCACTGCTCTGCATCTCCGAGGGTGT 600
Db 541 AGACCTTACTTTAGGAAGATGGAACACAGGATGCACTGCTCTGCATCTCCGAGGGTGT 600
QY 601 CCGGAGCCCACTGTGGAGTGGGTGCTCTGCACTCCCAAGGGAAGCTGTAAGAAGAA 660
Db 601 CCGGAGCCCACTGTGGAGTGGGTGCTCTGCACTCCCAAGGGAAGCTGTAAGAAGAA 660
QY 661 GGCCCTGCTGTTGTCAGAAAGAGGAAAGTACTTCAATGAGTGTTCGGAACAGACATC 720
Db 661 GGCCCTGCTGTTGTCAGAAAGAGGAAAGTACTTCAATGAGTGTTCGGAACAGACATC 720
QY 721 AGATGCTGTGCTAGAAATGCACTGGCGCGGAATGCCAACCAAGCTGTTCCACATAGATCTA 780
Db 721 AGATGCTGTGCTAGAAATGCACTGGCGCGGAATGCCAACCAAGCTGTTCCACATAGATCTA 780
QY 781 AACGAGCTCTCAGAGCACACTGCCCGCAGTTATTTCTGAAAGTGGGGGAAACCTTTGG 840
Db 781 AACGAGCTCTCAGAGCACACTGCCCGCAGTTATTTCTGAAAGTGGGGGAAACCTTTGG 840
QY 841 ATCAGGTGTAAGGCCATCCATGTGAACCATGGATTGCGGCTCACCTGGAGCTGGAAGAC 900
Db 841 ATCAGGTGTAAGGCCATCCATGTGAACCATGGATTGCGGCTCACCTGGAGCTGGAAGAC 900
QY 901 AAAGCCCTGGAGGAGGCGAGTACTTTGAGATGAGTACCTTCAAAAACAGAGACCATG 960
Db 901 AAAGCCCTGGAGGAGGCGAGTACTTTGAGATGAGTACCTTCAAAAACAGAGACCATG 960
QY 961 ATTCGGATTCTCTTGGCTTTGTTGTTCCGTTGGGAGGAACGACACCGGATATTACCC 1020
Db 961 ATTCGGATTCTCTTGGCTTTGTTGTTCCGTTGGGAGGAACGACACCGGATATTACCC 1020
QY 1021 TGCTCTTCTCAAAGCACCCCGAGCAGTCCAGCGTGTGGTGAACCATCTTAGAAAAAGGGTTT 1080
Db 1021 TGCTCTTCTCAAAGCACCCCGAGCAGTCCAGCGTGTGGTGAACCATCTTAGAAAAAGGGTTT 1080
QY 1081 ATAAAGCTTACAGCTCCGAGAGAGTATGAAATGACCCGTTACGAAAGTTCTGCTTC 1140
Db 1081 ATAAAGCTTACAGCTCCGAGAGAGTATGAAATGACCCGTTACGAAAGTTCTGCTTC 1140
QY 1141 TCAGTCAGGTTTAAAGCGTACCCAGATCCGATGCACTGAGTCTTCTCTCAAGCTCA 1200
Db 1141 TCAGTCAGGTTTAAAGCGTACCCAGATCCGATGCACTGAGTCTTCTCTCAAGCTCA 1200
QY 1201 TTTCTTGTGAACAGAGAGGCTGGAGGTGGGTACAGCATATCTAAATTTTTCGATCAT 1260

```

1201 TTTCTTGTGAACAGAGAGGCGCTGGAGGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260 Db  
1261 AAGAACAGCCAGGAGAGTACATATCTATGACAGAAATGATGACGCCCGAGTTCAACAAA 1320 Qy  
1261 AAGAACAGCCAGGAGAGTACATATCTATGACAGAAATGATGACGCCCGAGTTCAACAAA 1320 Db  
1321 ATGTTTCAGCTGAATATATAGAAAGAAACCTCAAGTGTCTAGCAAAATGCTCAGCCAGCCAG 1380 Qy  
1321 ATGTTTCAGCTGAATATATAGAAAGAAACCTCAAGTGTCTAGCAAAATGCTCAGCCAGCCAG 1380 Db  
1381 CGCTCTGTCTCTGTATGGCTACCCGCTACCCCTCTTGACCTTGGAAGAGTGTTCGGAC 1440 Qy  
1381 CGCTCTGTCTCTGTATGGCTACCCGCTACCCCTCTTGACCTTGGAAGAGTGTTCGGAC 1440 Db  
1441 AAATCTCCAAATGACAGGAGAAATCCCAAGAGAGTTTGAATAAAGAGCTAACAGA 1500 Qy  
1441 AAATCTCCAAATGACAGGAGAAATCCCAAGAGAGTTTGAATAAAGAGCTAACAGA 1500 Db  
1501 AAAGTGTGTTGGCCAGTGGGTGTCGAGCAGTACTCTAAATATAGTGAAGCCGCGGAAGGG 1560 Qy  
1501 AAAGTGTGTTGGCCAGTGGGTGTCGAGCAGTACTCTAAATATAGTGAAGCCGCGGAAGGG 1560 Db  
1561 CTTCCTGTCCTGTCTGTGCTGCTACAAATCTATGGGCAAGTCTTGGAAACCAATCTTTTAA 1620 Qy  
1561 CTTCCTGTCCTGTCTGTGCTGCTACAAATCTATGGGCAAGTCTTGGAAACCAATCTTTTAA 1620 Db  
1621 AACTCACAGAGCCCTTCCCTTTCTATGCAAGCAACATCTCTCTATGCGGACCATTTGGG 1680 Qy  
1621 AACTCACAGAGCCCTTCCCTTTCTATGCAAGCAACATCTCTCTATGCGGACCATTTGGG 1680 Db  
1681 CTCTCTCTCCCTTCAATGTTCTCTCATTTGTTGATCTGCCACAAATACAAAAGCAA 1740 Qy  
1681 CTCTCTCTCCCTTCAATGTTCTCTCATTTGTTGATCTGCCACAAATACAAAAGCAA 1740 Db  
1741 TTTAGTACAGAGTCAGCTGACAGATGATCAGGTGACTGGGCCCTCTGATACAGGATAC 1800 Qy  
1741 TTTAGTACAGAGTCAGCTGACAGATGATCAGGTGACTGGGCCCTCTGATACAGGATAC 1800 Db  
1801 TTTCTAGTTGACTTCAGGAGCTATGAATATGACCTTAAAGTGGAGTTCCCGAGAGAAAC 1860 Qy  
1801 TTTCTAGTTGACTTCAGGAGCTATGAATATGACCTTAAAGTGGAGTTCCCGAGAGAAAC 1860 Db  
1861 TTAGAGTTTGGAAAGTCTTGGGTTCTGGGCTTTCCGGAGGGTGAACGCCACCGCC 1920 Qy  
1861 TTAGAGTTTGGAAAGTCTTGGGTTCTGGGCTTTCCGGAGGGTGAACGCCACCGCC 1920 Db  
1921 TATGCAATTAGTAAACCGGAGTCTCAAATCAGTGGCGGTGAAGATGCTTAAAGAGAAA 1980 Qy  
1921 TATGCAATTAGTAAACCGGAGTCTCAAATCAGTGGCGGTGAAGATGCTTAAAGAGAAA 1980 Db  
1981 GCTGACAGCTGTGAAAAGAAAGCTCTCATGTGGAGCTCAAAATGATGACCCACCTGGGA 2040 Qy  
1981 GCTGACAGCTGTGAAAAGAAAGCTCTCATGTGGAGCTCAAAATGATGACCCACCTGGGA 2040 Db  
2041 GACCATGACAAATCGTGAATCTGCTGGGGGATGACACATGTCAGGGCCAGTGTACTTG 2100 Qy  
2041 GACCATGACAAATCGTGAATCTGCTGGGGGATGACACATGTCAGGGCCAGTGTACTTG 2100 Db  
2101 ATTTTGAATATTTGTTGATGTTGACCTCTCAATCTACCTAAGAGTAAAGAGAGAAAG 2160 Qy  
2101 ATTTTGAATATTTGTTGATGTTGACCTCTCAATCTACCTAAGAGTAAAGAGAGAAAG 2160 Db  
2161 TTTTCAGGACATGGAACAGATTTTTTAAGGAAACATAATTTTCAGTTCTTACCTACTTTC 2220 Qy  
2161 TTTTCAGGACATGGAACAGATTTTTTAAGGAAACATAATTTTCAGTTCTTACCTACTTTC 2220 Db  
2221 CAGGACATTTCAAATTCAGAGTCCTGTTTACGAGAGTTTCAGTTTACCCCGCCCTTG 2280 Qy  
2221 CAGGACATTTCAAATTCAGAGTCCTGTTTACGAGAGTTTCAGTTTACCCCGCCCTTG 2280 Db  
2281 GATCAGCTCTCAGGTTCAATGGGAATTCAAATTCATTTAGATGAGATGGAATGAA 2340 Qy  
2281 GATCAGCTCTCAGGTTCAATGGGAATTCAAATTCATTTAGATGAGATGGAATGAA 2340 Db

2341 AACCAAGAGGCTGCGCAAGAAAGAGAGGAAGATTTTCAAAGTGTGACGTTTGAAGAC 2400 Qy  
2341 AACCAAGAGGCTGCGCAAGAAAGAGAGGAAGATTTTCAAAGTGTGACGTTTGAAGAC 2400 Db  
2401 CTCTCTTGTGTTGCGTACCAAGTGGCCAAAGGCATGGAATCTCTGAGTTCAAAGTGTGT 2460 Qy  
2401 CTCTCTTGTGTTGCGTACCAAGTGGCCAAAGGCATGGAATCTCTGAGTTCAAAGTGTGT 2460 Db  
2461 GTCCACAGAGACCTGGCAGCCAGGATGTGTTGTCACCCACGGGAAGTGTGGAAGATC 2520 Qy  
2461 GTCCACAGAGACCTGGCAGCCAGGATGTGTTGTCACCCACGGGAAGTGTGGAAGATC 2520 Db  
2521 TGTGACTTTTGAAGTGGCCGAGACATCTCTGAGGACTCTCAGCTACCTGTCAGGGGCAAC 2580 Qy  
2521 TGTGACTTTTGAAGTGGCCGAGACATCTCTGAGGACTCTCAGCTACCTGTCAGGGGCAAC 2580 Db  
2581 GCACGGCTGCGCGTGAAGTGGATGGCACCCGAGAGCTTATTTGAAAGGATCTACACAATC 2640 Qy  
2581 GCACGGCTGCGCGTGAAGTGGATGGCACCCGAGAGCTTATTTGAAAGGATCTACACAATC 2640 Db  
2641 AAGAGTGAGCTGTGTCCTACCGCATCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700 Qy  
2641 AAGAGTGAGCTGTGTCCTACCGCATCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700 Db  
2701 CCTTACCTGTCATCTCTGTCGACGCTAACTTCTATAAACTGATTCAGAGTGAATTTAAA 2760 Qy  
2701 CCTTACCTGTCATCTCTGTCGACGCTAACTTCTATAAACTGATTCAGAGTGAATTTAAA 2760 Db  
2761 ATGGAGCAGCCATCTATGCCACAGAAAGGATATCTTGTAAATGCAATCTCTGCTGGCT 2820 Qy  
2761 ATGGAGCAGCCATCTATGCCACAGAAAGGATATCTTGTAAATGCAATCTCTGCTGGCT 2820 Db  
2821 TTTGACTCAAGGAAAGCGGCCATCTTCTCCCAACCTGACTTCAATTTTAGGATGTACAGCTG 2880 Qy  
2821 TTTGACTCAAGGAAAGCGGCCATCTTCTCCCAACCTGACTTCAATTTTAGGATGTACAGCTG 2880 Db  
2881 GCAGAGGCAAGAGAGCATGTATCAGAACATCCATCATCTACCAAAACAGGGCGGCCCT 2940 Qy  
2881 GCAGAGGCAAGAGAGCATGTATCAGAACATCCATCATCTACCAAAACAGGGCGGCCCT 2940 Db  
2941 CAGCAGAGAGGGGGCTCAGAGCCAGTCCACAGCGCCAGGTGAAGATTTCAACAGAAA 3000 Qy  
2941 CAGCAGAGAGGGGGCTCAGAGCCAGTCCACAGCGCCAGGTGAAGATTTCAACAGAAA 3000 Db  
3001 AGAAGTTAGCAGAGGAGCGCTTGGACCCGCCACCTAGCAGGCTGTAGACCGCAGAGCCA 3060 Qy  
3001 AGAAGTTAGCAGAGGAGCGCTTGGACCCGCCACCTAGCAGGCTGTAGACCGCAGAGCCA 3060 Db  
3061 AGATTAGCCTCGCTCTGAGGAAAGCGCCCTACAGCGGTTGCTTGGTGGACCTTTTCTCT 3120 Qy  
3061 AGATTAGCCTCGCTCTGAGGAAAGCGCCCTACAGCGGTTGCTTGGTGGACCTTTTCTCT 3120 Db  
3121 AGATGCTGTCTGCCATTAATCTCAAAAGTGAATCTTATAAACTCAAACTCTCTCGCACAG 3180 Qy  
3121 AGATGCTGTCTGCCATTAATCTCAAAAGTGAATCTTATAAACTCAAACTCTCTCGCACAG 3180 Db  
3181 GCGGAGAGCCAAATATGAGACTTTGTTGGTGGACCCGCTTACCTGGGGGCTTTTCCACG 3240 Qy  
3181 GCGGAGAGCCAAATATGAGACTTTGTTGGTGGACCCGCTTACCTGGGGGCTTTTCCACG 3240 Db  
3241 AGCTTGAAGGGAAGCCATGTATCTGAAATATAGTATATTTTGTAAATACGTGAAACAA 3300 Qy  
3241 AGCTTGAAGGGAAGCCATGTATCTGAAATATAGTATATTTTGTAAATACGTGAAACAA 3300 Db  
3301 ACCAAACCCGTTTTTGTCTAAAGGAGCTTAATATGATTTTAAAACTATGTTTTTAA 3360 Qy  
3301 ACCAAACCCGTTTTTGTCTAAAGGAGCTTAATATGATTTTAAAACTATGTTTTTAA 3360 Db  
3361 AATACATGCTAACTTTTCTATCTATTTAGTGTATATTTTATGATGGAATTAACCTTTC 3420 Qy  
3361 AATACATGCTAACTTTTCTATCTATTTAGTGTATATTTTATGATGGAATTAACCTTTC 3420 Db

```

Qy 3421 TACTGTAAAAA##### 3453
Db 3421 TACTGTAAAAA##### 3453

RESULT 5
US-07-906-397A-1
; Sequence 1, Application US/07906397A
; Patent No. 5621090
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMCLONE SYSTEMS INCORPORATED
; STREET: 180 VARICK STREET
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/906,397A
; FILING DATE: 19920626
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-PPPPPP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3453 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31..3009
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: 31..3006
; US-07-906-397A-1

Query Match 100.0%; Score 3453; DB 1; Length
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3453; Conservative 0; Mismatches 0; Indels
1 GCAGCGCTGGCTACCGCGCGCTCCGGAGGCCATGCGGCGTGTGGCGCA

```

Dd		1	GCGGCTGTGGCTACCGGCCGCCTCGSAGSGCCATCAGGGCGTTGGCGAGCGCACGCCG	60
Qy		61	CGGCTGTCTGCTTGTTTGTTCAGTAATCATCTTTGAGACCCTTACAACCAAGAC	120
Dd		61.	CGGCTGTCTGCTTGTTTGTTCAGTAATCATCTTTGAGACCCTTACAACCAAGAC	120.
Qy		121	CTGCCGTGATCAAGTGTGTTTTAATCAGTCAATGAACAAATGGCTCATCAGCGGAAAAG	180
Dd		121	CTGCCGTGATCAAAGTGTGTTTTAATCAGTCAATGAACAAATGGCTCATCAGCGGAAAAG	180.
Qy		181	CCATCATCTGATTACCAATGTCGAGATCCCAGAGACTTCCAAGTGTATCAGCGGAAAG	240
Dd		181	CCATCATCTGATTACCAATGTCGAGATCCCAGAGACTTCCAAGTGTATCAGCGGAAAG	240.
Qy		241	CAGAGTGAAAGGACGGGTATATGAAGCGGCCACCGTGGAGGTGGCCGAGTCTGGGTCCATC	300
Dd		241	CAGAGTGAAAGGACGGGTATATGAAGCGGCCACCGTGGAGGTGGCCGAGTCTGGGTCCATC	300.
Qy		301	ACCTGCAAGTGCAGCTCGCCA CCCAGGGGACCTTTCTTGCTCTGGGTCTTTTAAGCAC	360
Dd		301	ACCTGCAAGTGCAGCTCGCCA CCCAGGGGACCTTTCTTGCTCTGGGTCTTTTAAGCAC	360.
Qy		361	AGCTCCCTGGGCTGCCAGCCGACCTTTGATTTACAAAAACAGAGGAATCGTTTCCATGGCC	420
Dd		361	AGCTCCCTGGGCTGCCAGCCGACCTTTGATTTACAAAAACAGAGGAATCGTTTCCATGGCC	420.
Qy		421	AFTTTGAACGTGACAGAGACCAGGCAGAGAAATACCTACTCCATAATTACAGAGCGAACGC	480
Dd		421	AFTTTGAACGTGACAGAGACCAGGCAGAGAAATACCTACTCCATAATTACAGAGCGAACGC	480.
Qy		481	GCCAACTACA CAGTACTGTTCCAGTGAATGTAAAGAGATACACAGCTGTATGTGCTTAAGG	540
Dd		481	GCCAACTACA CAGTACTGTTCCAGTGAATGTAAAGAGATACACAGCTGTATGTGCTTAAGG	540.
Qy		541	AGACCTTACTTTAGGAAGATGGMAAAC CAGAGTGCATCTGCTCTGCATCTCCAGGGGTGT	600
Dd		541	AGACCTTACTTTAGGAAGATGGMAAAC CAGAGTGCATCTGCTCTGCATCTCCAGGGGTGT	600.
Qy		601	COGGAGCCCA CTGTGGAGTGGGTGCTCTGCA GCTCCCCACAGGAAAGCTGTAAAGAAGAA	660
Dd		601	COGGAGCCCA CTGTGGAGTGGGTGCTCTGCA GCTCCCCACAGGAAAGCTGTAAAGAAGAA	660.
Qy		661	GGCCCTGCTGTTGTACGAAGAGGAGAA AAGTACTTTCATGAGTTGTTCCGSAAACACATC	720
Dd		661	GGCCCTGCTGTTGTACGAAGAGGAGAA AAGTACTTTCATGAGTTGTTCCGSAAACACATC	720.
Qy		721	AGATGCTGTGTACGAATGC ACTGGSCCGCAATGC ACCAGCTGTTTCCATAGATCTTA	780
Dd		721	AGATGCTGTGTACGAATGC ACTGGSCCGCAATGC ACCAGCTGTTTCCATAGATCTTA	780.
Qy		781	AACAGGCTCCTCAGAGCACACTGCCCCAGTTATTCCTGAAAGTGGGGGAACCCCTGTGG	840
Dd		781	AACAGGCTCCTCAGAGCACACTGCCCCAGTTATTCCTGAAAGTGGGGGAACCCCTGTGG	840.
Qy		841	ATCAGTGTAAAGCCATCCATGTGAAC CATGGATTCCGGGCTCACTCGGAGCTGGAAGAC	900
Dd		841	ATCAGTGTAAAGCCATCCATGTGAAC CATGGATTCCGGGCTCACTCGGAGCTGGAAGAC	900.
Qy		901	AAAGCCCTGGAGGAGGCGACTATT TGATCAGTAGTACTTACCA CAACAGAGCACATG	960
Dd		901	AAAGCCCTGGAGGAGGCGACTATT TGATCAGTAGTACTTACCA CAACAGAGCACATG	960.
Qy		961	ATT CGGATTCCTT TGGCCTTT GTGTCTTCGTTGGAGGAA AGACACCGGATATTACCC	1020
Dd		961	ATT CGGATTCCTT TGGCCTTT GTGTCTTCGTTGGAGGAA AGACACCGGATATTACCC	1020.
Qy		1021	TGCTCTTCTCAAGACACCCAGCCAGTCAGCGTTGGTGA CCA TCCTAGAAAAAGGGTTT	1080
Dd		1021	TGCTCTTCTCAAGACACCCAGCCAGTCAGCGTTGGTGA CCA TCCTAGAAAAAGGGTTT	1080.
Qy		1081	ATAAACCGCTACCAAGCTCGAAGAAAGTAGTAATAATGA CCCGTACGAAAAGTTCTGCTTC	1140
Dd		1081	ATAAACCGCTACCAAGCTCGAAGAAAGTAGTAATAATGA CCCGTACGAAAAGTTCTGCTTC	1140.

1141 TCAGTCAGGTTTAAAGCGTACCCAGCAATCCGATCGACGTGGATCTTCTCTCAAGCCTCA 1200  
1141 TCAGTCAGGTTTAAAGCGTACCCAGCAATCCGATCGACGTGGATCTTCTCTCAAGCCTCA 1200  
1201 TTTCTCTGTGAAACAGAGAGGCGCTGGAGGATGGGTACAGCATATCTAAATTTTGGATCAT 1260  
1201 TTTCTCTGTGAAACAGAGAGGCGCTGGAGGATGGGTACAGCATATCTAAATTTTGGATCAT 1260  
1261 AAGAAACAAGCCAGGAGATGATATCTATGTCAGAAAATGATGACGCCAGTTTCAACAAA 1320  
1261 AAGAAACAAGCCAGGAGATGATATCTATGTCAGAAAATGATGACGCCAGTTTCAACAAA 1320  
1321 ATGTTCACGCTGAATATAGAAAGAAACCTCAAGTGTCTAGCAATGCTCAGCCAGCCAG 1380  
1321 ATGTTCACGCTGAATATAGAAAGAAACCTCAAGTGTCTAGCAATGCTCAGCCAGCCAG 1380  
1381 GCGTCTCTGTCTCTGATGGCTACCCGCTACCTCTTTGGACCTGGAAGAAGTGTTCGGAC 1440  
1381 GCGTCTCTGTCTCTGATGGCTACCCGCTACCTCTTTGGACCTGGAAGAAGTGTTCGGAC 1440  
1441 AAATCTCCCAATTTGACGGAGGAAATCCCAAGAGGATTTGGAAATAAAAAGGCTAACAGA 1500  
1441 AAATCTCCCAATTTGACGGAGGAAATCCCAAGAGGATTTGGAAATAAAAAGGCTAACAGA 1500  
1501 AAGTGTGTTGGCCAGTGGGTGTCAGCAGTACTCTAAATATGATGAGGCGCGGAAGGG 1560  
1501 AAGTGTGTTGGCCAGTGGGTGTCAGCAGTACTCTAAATATGATGAGGCGCGGAAGGG 1560  
1561 CTTCTGGTCAAAATGCTGTGCGTACAAATTTCTATGGGCACGCTCTTGGCAAAACATCTTTTAA 1620  
1561 CTTCTGGTCAAAATGCTGTGCGTACAAATTTCTATGGGCACGCTCTTGGCAAAACATCTTTTAA 1620  
1621 AACTACCAAGGCGCCCTTCCCTTTTCATCCAAAGCAACATCTCTCTATGCGACCAATTGGG 1680  
1621 AACTACCAAGGCGCCCTTCCCTTTTCATCCAAAGCAACATCTCTCTATGCGACCAATTGGG 1680  
1681 CTTCTGCTCCCTTCAATGTTGTTCTCATTTGTTGATCTGCCACAAATACAAAAGCAA 1740  
1681 CTTCTGCTCCCTTCAATGTTGTTCTCATTTGTTGATCTGCCACAAATACAAAAGCAA 1740  
1741 TTTAGTACGAGAGTCAGCTGACAGATGATCCAGGTGATCTGGGCCCTGGATACAGAGTAC 1800  
1741 TTTAGTACGAGAGTCAGCTGACAGATGATCCAGGTGATCTGGGCCCTGGATACAGAGTAC 1800  
1801 TTCTACGTTGACTTCAGGGAATATGAATATGACCTTAAGTGGGAGTTCCCGAGAGAAAC 1860  
1801 TTCTACGTTGACTTCAGGGAATATGAATATGACCTTAAGTGGGAGTTCCCGAGAGAAAC 1860  
1861 TTAGAGTTTGGAAAGTCTGGGCTCTGGGCTTTCCGGAGGGTATGAACGCCAGGCC 1920  
1861 TTAGAGTTTGGAAAGTCTGGGCTCTGGGCTTTCCGGAGGGTATGAACGCCAGGCC 1920  
1921 TATGGCATTTAGTAAACGGAGTCTCAATTCAGGTGGCGTGAAGATGCTTAAAGAGAAA 1980  
1921 TATGGCATTTAGTAAACGGAGTCTCAATTCAGGTGGCGTGAAGATGCTTAAAGAGAAA 1980  
1981 GCTGACAGCTGTGAAAAGAAAGCTCTCATGTCCGAGCTCAAAATGATGACCCACCTGGGA 2040  
1981 GCTGACAGCTGTGAAAAGAAAGCTCTCATGTCCGAGCTCAAAATGATGACCCACCTGGGA 2040  
2041 CACCATGACAAATCGTGAATCTGTGGGGGATGACATCTGTACGGGCGAGTGTACTTGG 2100  
2041 CACCATGACAAATCGTGAATCTGTGGGGGATGACATCTGTACGGGCGAGTGTACTTGG 2100  
2101 ATTTTGTGAATATGTTGTCTATGCTGACCTTCACTCACTTAAGAGTAAAGAGAGAG 2160  
2101 ATTTTGTGAATATGTTGTCTATGCTGACCTTCACTCACTTAAGAGTAAAGAGAGAG 2160  
2161 TTTTCAAGGACATGACAGAGATTTTAAAGGAACATAATTTTCACTTACCTACTTTTC 2220  
2161 TTTTCAAGGACATGACAGAGATTTTAAAGGAACATAATTTTCACTTACCTACTTTTC 2220

2221 CAGGCACATTTCAAATTCAGCATGCCCTGGTTTACGAGAAGTTTCAAGTTACACCGCCCTTG 2280  
2221 CAGGCACATTTCAAATTCAGCATGCCCTGGTTTACGAGAAGTTTCAAGTTACACCGCCCTTG 2280  
2281 GATCAGCTCTCAGGTTTCAATGGGAAATTCATTTCAATCTGGAAGATGAGATTAATGAA 2340  
2281 GATCAGCTCTCAGGTTTCAATGGGAAATTCATTTCAATCTGGAAGATGAGATTAATGAA 2340  
2341 AACCAAGAGGCTGGCAGAGAAGAGAGAGAGATTTTGAACGTGCTGACGTTTGAAGAC 2400  
2341 AACCAAGAGGCTGGCAGAGAAGAGAGAGAGATTTTGAACGTGCTGACGTTTGAAGAC 2400  
2401 CTCCTTTGCTTGTCTGATACCAAGTGGCCAAAGGCAATGAAATTTCTTGGAGTTCAAGTCTG 2460  
2401 CTCCTTTGCTTGTCTGATACCAAGTGGCCAAAGGCAATGAAATTTCTTGGAGTTCAAGTCTG 2460  
2461 GTCCAAGAGACCTGGCAGCCAGCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2520  
2461 GTCCAAGAGACCTGGCAGCCAGCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2520  
2521 TGTGACTTTTGGACTGGCCCGAGACATCTCTGAGCGACTCCAGCTACGTCGTGAGGGGCAAC 2580  
2521 TGTGACTTTTGGACTGGCCCGAGACATCTCTGAGCGACTCCAGCTACGTCGTGAGGGGCAAC 2580  
2581 GCACGCTGCGGTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGGATCTACACAATC 2640  
2581 GCACGCTGCGGTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGGATCTACACAATC 2640  
2641 AAGAGTGAACCTCTGCTGTCTACGGCATCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700  
2641 AAGAGTGAACCTCTGCTGTCTACGGCATCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700  
2701 CTTTACCTCTGGCAATCTCTGTGACGCTTAATCTTATAAACTGATTCAGAGTGGATTTAAA 2760  
2701 CTTTACCTCTGGCAATCTCTGTGACGCTTAATCTTATAAACTGATTCAGAGTGGATTTAAA 2760  
2761 ATGAGCAGGCAATCTATGCGCACAGAGGATATATCTTTGTAATGCAATCTCTGCTGGCT 2820  
2761 ATGAGCAGGCAATCTATGCGCACAGAGGATATATCTTTGTAATGCAATCTCTGCTGGCT 2820  
2821 TTTGACTCAAGGAAGCGGCCATCTTCCCAACCTGACTTCTATTTTAGATGTCAGCTG 2880  
2821 TTTGACTCAAGGAAGCGGCCATCTTCCCAACCTGACTTCTATTTTAGATGTCAGCTG 2880  
2881 GCAGAGCGAAGAGCATGTATCAGAACTCCATCCATCTACCAAAACAGGCGGCCCT 2940  
2881 GCAGAGCGAAGAGCATGTATCAGAACTCCATCCATCTACCAAAACAGGCGGCCCT 2940  
2941 CAGCAGAGGCGGCTCAGAGCCCATGTGCCACAGCGCCAGGTGAAGATTCACAGAGAA 3000  
2941 CAGCAGAGGCGGCTCAGAGCCCATGTGCCACAGCGCCAGGTGAAGATTCACAGAGAA 3000  
3001 AGAAGTTTACGAGGAGGCTTGGACCCCGCCACCTAGCAGGCTGTAGACCGCAGAGCCA 3060  
3001 AGAAGTTTACGAGGAGGCTTGGACCCCGCCACCTAGCAGGCTGTAGACCGCAGAGCCA 3060  
3061 AGATTAGCCTCGCTCTGAGGAAGCGCCCTACAGCGCTTCTGCTGGAATTTTCTCT 3120  
3061 AGATTAGCCTCGCTCTGAGGAAGCGCCCTACAGCGCTTCTGCTGGAATTTTCTCTCT 3120  
3121 AGATGCTGTCTGCATTAATCTCCAAAGTGAATCTTATAAATCAAACTCTCTCTGCAAG 3180  
3121 AGATGCTGTCTGCATTAATCTCCAAAGTGAATCTTATAAATCAAACTCTCTCTGCAAG 3180  
3181 GCGGAGAGCCCAATATGAGATCTTGTGTGAGCCCGCTACCTCGGGGCTTTTCCACG 3240  
3181 GCGGAGAGCCCAATATGAGATCTTGTGTGAGCCCGCTACCTCGGGGCTTTTCCACG 3240  
3241 AGCTTGAGGGGAAAGCCATGATCTGAAATATAGTATATTTCTGTAAATACGTGAACAA 3300  
3241 AGCTTGAGGGGAAAGCCATGATCTGAAATATAGTATATTTCTGTAAATACGTGAACAA 3300  
3301 ACCAAAACCGTTTTTGTCTAAGGAAAGCTAAATATGATTTTTTAAAAATCTATGTTTAA 3360





Db 721 AGATGCTGTCTAGAAATGCACTGGGCGCGAATGACCAAGCTGTTTCAACCATAGATCTA 780  
Qy  
781 AACCAAGCTCTCAGAGACACACTGCCCCAGTATTCTCTGAAAGTGGGGGAACCCCTTGTG 840  
Db  
781 AACCAAGCTCTCAGAGACACACTGCCCCAGTATTCTCTGAAAGTGGGGGAACCCCTTGTG 840  
Qy  
841 ATCAGGTGTAAAGGCATTCATGTGAAACCATGTGATTTCGGGTCTACCTCGGAGCTGGAAGAC 900  
Db  
841 ATCAGGTGTAAAGGCATTCATGTGAAACCATGTGATTTCGGGTCTACCTCGGAGCTGGAAGAC 900  
Qy  
901 AAAGCCCTGGAGGAGGCGACCTACTTTGAGATGAGTACTCTCCACAAACAGGACCATG 960  
Db  
901 AAAGCCCTGGAGGAGGCGACCTACTTTGAGATGAGTACTCTCCACAAACAGGACCATG 960  
Qy  
961 ATTCCGATTCCTTTGGCTTTGTCTTCCTCGTGGGAAGGAACGACCGGATATTACACC 1020  
Db  
961 ATTCCGATTCCTTTGGCTTTGTCTTCCTCGTGGGAAGGAACGACCGGATATTACACC 1020  
Qy  
1021 TGCTCTTCCTCAAGACACCCAGCCAGTCAGCGTTGGTGACCTCTAGAAAAAGGGTTT 1080  
Db  
1021 TGCTCTTCCTCAAGACACCCAGCCAGTCAGCGTTGGTGACCTCTAGAAAAAGGGTTT 1080  
Qy  
1081 ATAAACGCTACCAAGCTCGCAAGAGAGTATGAAATTGACCCGTCAGAAAAAGTTCTGCTTC 1140  
Db  
1081 ATAAACGCTACCAAGCTCGCAAGAGAGTATGAAATTGACCCGTCAGAAAAAGTTCTGCTTC 1140  
Qy  
1141 TCAGTCAGGTTTAAAGCGTACCAACGAATCCGATGCACTGGATCTCTCTCAAGCCTCA 1200  
Db  
1141 TCAGTCAGGTTTAAAGCGTACCAACGAATCCGATGCACTGGATCTCTCTCAAGCCTCA 1200  
Qy  
1201 TTTCTCTGTGAACAGAGAGCGCTGGAGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260  
Db  
1201 TTTCTCTGTGAACAGAGAGCGCTGGAGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260  
Qy  
1261 AAGAAACAGCCAGGAGAGTACATATTCTATGCAGAAAAATGATGACGCCAGTTCCACAAA 1320  
Db  
1261 AAGAAACAGCCAGGAGAGTACATATTCTATGCAGAAAAATGATGACGCCAGTTCCACAAA 1320  
Qy  
1321 ATGTTCAAGCTGAATATAAGAAAGAAACCTCAAGTGTCTAGCAAAATGCTCAGCCAGCAG 1380  
Db  
1321 ATGTTCAAGCTGAATATAAGAAAGAAACCTCAAGTGTCTAGCAAAATGCTCAGCCAGCAG 1380  
Qy  
1381 GCGTCTGTCTCTGATGCTACCCGCTACCCCTCTTGGACCTGGAAGAGTGTTCGGAC 1440  
Db  
1381 GCGTCTGTCTCTGATGCTACCCGCTACCCCTCTTGGACCTGGAAGAGTGTTCGGAC 1440  
Qy  
1441 AAATCTCCCAATTTGCACGGAGGAATCCCAAGAGAGTGTGGAAATAAAAGGCTTACAGA 1500  
Db  
1441 AAATCTCCCAATTTGCACGGAGGAATCCCAAGAGAGTGTGGAAATAAAAGGCTTACAGA 1500  
Qy  
1501 AAAGTGTTTGGCCAGTGGGTGTGAGCAGTACTCTAAATATGAGTGAGGCGGGAAGGG 1560  
Db  
1501 AAAGTGTTTGGCCAGTGGGTGTGAGCAGTACTCTAAATATGAGTGAGGCGGGAAGGG 1560  
Qy  
1561 CTTCTGGTCAAAATGCTGTGGTCAATCTATGGGCAAGCTTTCGGAACCAATCTTTTAA 1620  
Db  
1561 CTTCTGGTCAAAATGCTGTGGTCAATCTATGGGCAAGCTTTCGGAACCAATCTTTTAA 1620  
Qy  
1621 AACTCACCGGCCCCCTTCCTTCATCCAGACAACATCTCTCTATGGACCAATTTGGG 1680  
Db  
1621 AACTCACCGGCCCCCTTCCTTCATCCAGACAACATCTCTCTATGGACCAATTTGGG 1680  
Qy  
1681 CTCTGTCTCCCTTCATTTGTTGTTCTCAATTTGTTGATCTGCCACAAATACAAAAGCAA 1740  
Db  
1681 CTCTGTCTCCCTTCATTTGTTGTTCTCAATTTGTTGATCTGCCACAAATACAAAAGCAA 1740  
Qy  
1741 TTTAGGTACAGAGTCAAGTGTGATGATCCAGGTGATCTGGCCCCCTTGGATTAACAGTAC 1800  
Db  
1741 TTTAGGTACAGAGTCAAGTGTGATGATCCAGGTGATCTGGCCCCCTTGGATTAACAGTAC 1800  
Qy  
1801 TTTCTAGTTTCACTCAGGACTATGATATGACCTTAAGTGGAGTTCCCGAGAGAGAC 1860  
Db  
1801 TTTCTAGTTTCACTCAGGACTATGATATGACCTTAAGTGGAGTTCCCGAGAGAGAC 1860

Qy 1861 TTAGAGTTTCGGAGAGTCTCTGGGCTCTGGCGCTTTTCGGAGGGGTGATGAACCCACCGCC 1920  
Db 1861 TTAGAGTTTCGGAGAGTCTCTGGGCTCTGGCGCTTTTCGGAGGGGTGATGAACCCACCGCC 1920  
Qy 1921 TATGGCAATTAGTATAAACCGGAGTCTCAATTCAAGTGGCGGTGAAGATGCTTAAAAAGAGAA 1980  
Db 1921 TATGGCAATTAGTATAAACCGGAGTCTCAATTCAAGTGGCGGTGAAGATGCTTAAAAAGAGAA 1980  
Qy 1981 GCTGACAGCTGTGAAAAAAGAGCTCTCATGTGGAGCTCAAAATGATGACCCACCTGGGA 2040  
Db 1981 GCTGACAGCTGTGAAAAAAGAGCTCTCATGTGGAGCTCAAAATGATGACCCACCTGGGA 2040  
Qy 2041 CACATGACAACATCGTGAATCTGCTGGGGGCAATGACACATGTCAGGGCCAGTGTACTTG 2100  
Db 2041 CACATGACAACATCGTGAATCTGCTGGGGGCAATGACACATGTCAGGGCCAGTGTACTTG 2100  
Qy 2101 ATTTTTCGAATATTGTTCTATGCTGCTCTCAACTCACTAAAGAAATGATAAGAGAGAG 2160  
Db 2101 ATTTTTCGAATATTGTTCTATGCTGCTCTCAACTCACTAAAGAAATGATAAGAGAGAG 2160  
Qy 2161 TTTTACAGGACATGACAGAGATTTTAAAGAAATATTTTCACTTCACTTCACTTCTTTC 2220  
Db 2161 TTTTACAGGACATGACAGAGATTTTAAAGAAATATTTTCACTTCACTTCACTTCTTTC 2220  
Qy 2221 CAGGCAATTTCAATTTCCAGATGCTGCTGTTTCAAGAGTTCAGTTTACACCCGCCCTTG 2280  
Db 2221 CAGGCAATTTCAATTTCCAGATGCTGCTGTTTCAAGAGTTCAGTTTACACCCGCCCTTG 2280  
Qy 2281 GATCAGCTCTCAGGTTTCAATTTGGAATTTCAATTTCAATTTCTGAAGATGAGATTGAATGAA 2340  
Db 2281 GATCAGCTCTCAGGTTTCAATTTGGAATTTCAATTTCAATTTCTGAAGATGAGATTGAATGAA 2340  
Qy 2341 AACAGAGAGGCTGGCAG 2400  
Db 2341 AACAGAGAGGCTGGCAG 2400  
Qy 2401 CTCCTTTGCTTTGCTTACAGTGGCCAAAGGATGGAATTTCTTGGAGTTCAGTCTGTGT 2460  
Db 2401 CTCCTTTGCTTTGCTTACAGTGGCCAAAGGATGGAATTTCTTGGAGTTCAGTCTGTGT 2460  
Qy 2461 GTCCACAGAGACCTGGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520  
Db 2461 GTCCACAGAGACCTGGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520  
Qy 2521 TGTGACTTTGGACTTGGCCCGGAGACATCTGAGCGACTCCAGCTACGTCAGTCTGAGGCGAAC 2580  
Db 2521 TGTGACTTTGGACTTGGCCCGGAGACATCTGAGCGACTCCAGCTACGTCAGTCTGAGGCGAAC 2580  
Qy 2581 GCACGGCTGCCGTTGAAAGTGGATGGCACCAGAGAGCTTATTTGAAGGGATCTTACAAATC 2640  
Db 2581 GCACGGCTGCCGTTGAAAGTGGATGGCACCAGAGAGCTTATTTGAAGGGATCTTACAAATC 2640  
Qy 2641 AAGAGTCAAGCTTGGTCTTACCGCATCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700  
Db 2641 AAGAGTCAAGCTTGGTCTTACCGCATCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700  
Qy 2701 CTTTACCTCTGSCATCTCTGTCGACGCTAACTTCTATAAACTGATTTCAAGTGGATTTAAA 2760  
Db 2701 CTTTACCTCTGSCATCTCTGTCGACGCTAACTTCTATAAACTGATTTCAAGTGGATTTAAA 2760  
Qy 2761 ATGAGCAGCCATTTCTATGCCACAGAGGGATATATCTTTGTAATGCAATCTCTGCTGGGCT 2820  
Db 2761 ATGAGCAGCCATTTCTATGCCACAGAGGGATATATCTTTGTAATGCAATCTCTGCTGGGCT 2820  
Qy 2821 TTTGACTCAAGAGAGCGGCCATCTTCCCAACCTGACTTCAATTTTAGGATGTGAGCTG 2880  
Db 2821 TTTGACTCAAGAGAGCGGCCATCTTCCCAACCTGACTTCAATTTTAGGATGTGAGCTG 2880  
Qy 2881 GCAGAGCAGAGAGAGATGATTCAGAACTTCACTTACCAAAACAGCGGCCCTCT 2940  
Db 2881 GCAGAGCAGAGAGAGATGATTCAGAACTTCACTTACCAAAACAGCGGCCCTCT 2940

QY 2941 CAGCAGAGAGGCGGCTCAGAGCCCGAGTCGCCACAGCGCCAGGTGAAGATTACAGAGAA 3000  
Db |||||  
QY 2941 CAGCAGAGAGGCGGCTCAGAGCCCGAGTCGCCACAGCGCCAGGTGAAGATTACAGAGAA 3000  
Db |||||  
QY 3001 AGAAGTTAGCAGAGGCGCTGGACCCCGCCACCTAGAGGCTGTAGACCGCAGAGCCA 3060  
Db |||||  
QY 3001 AGAAGTTAGCAGAGGCGCTGGACCCCGCCACCTAGAGGCTGTAGACCGCAGAGCCA 3060  
Db |||||  
QY 3061 AGATTAGCCTCGCTCAGAGAGCGCCCTACAGCGGCTTGGCTGGAGCTTTTCT 3120  
Db |||||  
QY 3061 AGATTAGCCTCGCTCAGAGAGCGCCCTACAGCGGCTTGGCTGGAGCTTTTCT 3120  
Db |||||  
QY 3121 AGATGCTGTCTGCCATTACTCCAAAGTGACTTCTATAAAATCAAACTCTCTCCACAG 3180  
Db |||||  
QY 3121 AGATGCTGTCTGCCATTACTCCAAAGTGACTTCTATAAAATCAAACTCTCTCCACAG 3180  
Db |||||  
QY 3181 GCGGAGAGGCAATAATAGACTTGTGGTGGAGCGCCCTACCTGGGGCTTTTCCACG 3240  
Db |||||  
QY 3181 GCGGAGAGGCAATAATAGACTTGTGGTGGAGCGCCCTACCTGGGGCTTTTCCACG 3240  
Db |||||  
QY 3241 AGCTTGAGGGAAGCCATGATCTGAATATAGTATATTTCTTAATACGTGAACAA 3300  
Db |||||  
QY 3241 AGCTTGAGGGAAGCCATGATCTGAATATAGTATATTTCTTAATACGTGAACAA 3300  
Db |||||  
QY 3301 ACCAAACCCGTTTTTGTGAAGGAAAGCTAAATATGATTTTTTAAATCTATGTTTAA 3360  
Db |||||  
QY 3301 ACCAAACCCGTTTTTGTGAAGGAAAGCTAAATATGATTTTTTAAATCTATGTTTAA 3360  
Db |||||  
QY 3361 AATACTATGTAACCTTTTTCATCTATTAGTGATATATTTATGATGGAATAAATCTTC 3420  
Db |||||  
QY 3361 AATACTATGTAACCTTTTTCATCTATTAGTGATATATTTATGATGGAATAAATCTTC 3420  
Db |||||  
QY 3421 TACTGTAAAT 3453  
Db |||||  
QY 3421 TACTGTAAAT 3453  
Db |||||

## RESULT 7

US-09-021-324-1  
; Sequence 1, Application US/09021324  
; Patent No. 5912133  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Imclone Systems Incorporated  
; STREET: 180 Varick Street  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/021,324  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/977,451  
; FILING DATE: 1992-11-19  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/906,397  
; FILING DATE: 26-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US PCT/US92/05401  
; FILING DATE: 26-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: TW 81102961

; FILING DATE: 15-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US PCT/US92/02750  
; FILING DATE: 02-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/813,593  
; FILING DATE: 24-DEC-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/793,065  
; FILING DATE: 15-NOV-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/728,913  
; FILING DATE: 28-JUN-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/679,666  
; FILING DATE: 02-APR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Feit, Irving N.  
; REGISTRATION NUMBER: 28,601  
; REFERENCE/DOCKET NUMBER: LEM-3-7P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-645-1405  
; TELEFAX: 212-645-2054  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3453 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
; FEATURE:  
; NAME/KEY: mat\_peptide  
; LOCATION: 112..3006  
; FEATURE:  
; NAME/KEY: sig\_peptide  
; LOCATION: 31..111  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 31..3009  
; US-09-021-324-1  
Query Match 100.0%; Score 3453; DB 2; Length 3453;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCGGCGCTGCTACCGCGCGCTCCGAGAGGCGCATCGCGGCGTTGGCGCAGCGCAGCGCGG 60  
Db |||||  
QY 61 GCGGCGCTGCTACCGCGCGCTCCGAGAGGCGCATCGCGGCGTTGGCGCAGCGCAGCGCGG 60  
Db |||||  
QY 61 GCGGCGCTGCTACCGCGCGCTCCGAGAGGCGCATCGCGGCGTTGGCGCAGCGCAGCGCGG 60  
Db |||||  
QY 121 CTGCGCTGTGATCAAGTGCTGTTTAAATCAGTCATGAGCAATGGCTCATCAGCGGGAAG 180  
Db |||||  
QY 121 CTGCGCTGTGATCAAGTGCTGTTTAAATCAGTCATGAGCAATGGCTCATCAGCGGGAAG 180  
Db |||||  
QY 181 CCATCATCGTACCGAAATGGTGGAGGATCCCGAGAGACCTCCAGTGATACCCGAGGCGC 240  
Db |||||  
QY 181 CCATCATCGTACCGAAATGGTGGAGGATCCCGAGAGACCTCCAGTGATACCCGAGGCGC 240  
Db |||||  
QY 241 CAGAGTGAAGGACCGGTATATGAAGCGGCCACCGTGGAGGTGGCGAGTCTGGGTCCATC 300  
Db |||||  
QY 241 CAGAGTGAAGGACCGGTATATGAAGCGGCCACCGTGGAGGTGGCGAGTCTGGGTCCATC 300  
Db |||||  
QY 301 ACCCTGCAAGTGCAAGTCTCGCCACCCGAGGACCTTTCTGCTCTGGGTCTTTAAGCAC 360  
Db |||||  
QY 301 ACCCTGCAAGTGCAAGTCTCGCCACCCGAGGACCTTTCTGCTCTGGGTCTTTAAGCAC 360  
Db |||||  
QY 361 AGCTCCCTGGGCTGCGCAGCGCGCACTTTGATTTTACAAAACAGAGGAATCGTTTCCATGCC 420  
Db |||||

Db 361 AGCTCCCTGGGCTGGCCAGCCGCACTTTGATTTTACAAAAGAGGAATCGTTTCATGGCC 420  
QY 421 ATCTTGAACGTGACAGAGACCCAGGCGAGAGATACCTACTCCATATTCAGAGCGAACGC 480  
Db 421 ATCTTGAACGTGACAGAGACCCAGGCGAGAGATACCTACTCCATATTCAGAGCGAACGC 480  
QY 481 GCCAACTACAGACTACTGTTTCACAGTGAATGTAAAGATACACAGCTGTATGCTAAGG 540  
Db 481 GCCAACTACAGACTACTGTTTCACAGTGAATGTAAAGATACACAGCTGTATGCTAAGG 540  
QY 541 AGACCTTACTTTAGGAGATGGAACCCAGATGCACTGCTCTGCATCTCCGAGGGTGT 600  
Db 541 AGACCTTACTTTAGGAGATGGAACCCAGATGCACTGCTCTGCATCTCCGAGGGTGT 600  
QY 601 CCGGAGCCCACTGTGGAGTGGGTCTCTGCAGCTCCCAAGGGAAGCTGTAAAGAGAA 660  
Db 601 CCGGAGCCCACTGTGGAGTGGGTCTCTGCAGCTCCCAAGGGAAGCTGTAAAGAGAA 660  
QY 661 GGCCCTGCTGTTCAGAAAAGGAGGAAAAAGTACTTCATGAGTTGTTTCGGAAACAGACATC 720  
Db 661 GGCCCTGCTGTTCAGAAAAGGAGGAAAAAGTACTTCATGAGTTGTTTCGGAAACAGACATC 720  
QY 721 AGATGCTGTGTAGAAATGCACTGGGCGCGCAATGCCAACCAAGCTTTCACCATAGATCTA 780  
Db 721 AGATGCTGTGTAGAAATGCACTGGGCGCGCAATGCCAACCAAGCTTTCACCATAGATCTA 780  
QY 781 AACAGAGCTCTCAGAGACACACTGCCCGAGTATTCTTGAAGTGGGGGAACCCCTGTGG 840  
Db 781 AACAGAGCTCTCAGAGACACACTGCCCGAGTATTCTTGAAGTGGGGGAACCCCTGTGG 840  
QY 841 ATCAGGTGTAAAGCCATCCATGTGAAACCATGATTCGGGCTCAGCTGGAGCTGGAAGAC 900  
Db 841 ATCAGGTGTAAAGCCATCCATGTGAAACCATGATTCGGGCTCAGCTGGAGCTGGAAGAC 900  
QY 901 AAAGCCCTGGAGGGGAGGAGTACTTTGAGATGAGTACTTCCACAAAAGGAGCCATG 960  
Db 901 AAAGCCCTGGAGGGGAGGAGTACTTTGAGATGAGTACTTCCACAAAAGGAGCCATG 960  
QY 961 ATTCGGATTCCTTGGCCCTTGTGTCTTCCTGGGAGGAAGACGACACCGGATATTACACC 1020  
Db 961 ATTCGGATTCCTTGGCCCTTGTGTCTTCCTGGGAGGAAGACGACACCGGATATTACACC 1020  
QY 1021 TGCTCTTCCTCAAGACACCCAGCAGTCAGCGTTGGTGACCATCTAGAAAAAGGGTTT 1080  
Db 1021 TGCTCTTCCTCAAGACACCCAGCAGTCAGCGTTGGTGACCATCTAGAAAAAGGGTTT 1080  
QY 1081 ATAAACGGTACAGCTCGCAAGAGAGTATGAAATTGACCCGTAGCAAAAGTTCTGCTTC 1140  
Db 1081 ATAAACGGTACAGCTCGCAAGAGAGTATGAAATTGACCCGTAGCAAAAGTTCTGCTTC 1140  
QY 1141 TCAGTCAGGTTTAAAGCGTACCCAGGATCCGATGCACTGGATCTTCTCAAGCCCTCA 1200  
Db 1141 TCAGTCAGGTTTAAAGCGTACCCAGGATCCGATGCACTGGATCTTCTCAAGCCCTCA 1200  
QY 1201 TTTCTCTGTAACAGAGAGGCTGGAGGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260  
Db 1201 TTTCTCTGTAACAGAGAGGCTGGAGGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260  
QY 1261 AAGAACAAAGCCAGGAGAGTACATATTTCTATGAGAAAAATGATGACGCCAGTTTCAACAA 1320  
Db 1261 AAGAACAAAGCCAGGAGAGTACATATTTCTATGAGAAAAATGATGACGCCAGTTTCAACAA 1320  
QY 1321 ATGTTTCAGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCTTCAGCCAGCCAG 1380  
Db 1321 ATGTTTCAGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCTTCAGCCAGCCAG 1380  
QY 1381 GGGTCCCTGCTCTGATGGCTTACCGCTACCTCTTGGACCTGGAAGAGTGTTCGGAC 1440  
Db 1381 GGGTCCCTGCTCTGATGGCTTACCGCTACCTCTTGGACCTGGAAGAGTGTTCGGAC 1440  
QY 1441 AAATCTCCCAATTGCAACGGAGGAAATCCACAGAGGAGTTTGGAAATAAAGGCTTAACAGA 1500

1441 AAATCTCCCAATTGCAACGGAGGAAATCCACAGAGGAGTTTGGAAATAAAGGCTTAACAGA 1500  
QY 1501 AAAGTGTTTGGCCAGTGGGTGCGAGCAGTACTCTAAATATATGATGAGGCGCGGAAGGG 1560  
Db 1501 AAAGTGTTTGGCCAGTGGGTGCGAGCAGTACTCTAAATATATGATGAGGCGCGGAAGGG 1560  
QY 1561 CTTCTCGTCAAAATGCTGTGCTACAAATTTCTATGGGCACTGCTTTCGGGAAACATCTTTTAA 1620  
Db 1561 CTTCTCGTCAAAATGCTGTGCTACAAATTTCTATGGGCACTGCTTTCGGGAAACATCTTTTAA 1620  
QY 1621 AACTCACAGGCCCTTCCCTTTTCATCCAAAGACAAATCTCTCTCTATGAGCAATTTGGG 1680  
Db 1621 AACTCACAGGCCCTTCCCTTTTCATCCAAAGACAAATCTCTCTCTATGAGCAATTTGGG 1680  
QY 1681 CTCTGCTCCCTTTCATTTGTTGTTCTCATTTGTTGATCTGCGCAAAATACAAAGGCA 1740  
Db 1681 CTCTGCTCCCTTTCATTTGTTGTTGTTGATCTGCGCAAAATACAAAGGCA 1740  
QY 1741 TTTAGGTACAGAGTCACTGTCAGATGATCCAGTGTACTGGCCCTTGGGATAACAGATAC 1800  
Db 1741 TTTAGGTACAGAGTCACTGTCAGATGATCCAGTGTACTGGCCCTTGGGATAACAGATAC 1800  
QY 1801 TTTCTACGTTTTCAGGAGCTATGATATGACCTTAAAGTGGGAGTTTCCCGAGAGAGAAC 1860  
Db 1801 TTTCTACGTTTTCAGGAGCTATGATATGACCTTAAAGTGGGAGTTTCCCGAGAGAGAAC 1860  
QY 1861 TTTAGAGTTTGGGAGGTTCTGGGGTCTGGGGCTTTTCGGGAGGTTGATGAACGCCACGGCC 1920  
Db 1861 TTTAGAGTTTGGGAGGTTCTGGGGTCTGGGGCTTTTCGGGAGGTTGATGAACGCCACGGCC 1920  
QY 1921 TATGGCAATAGTAAACCGGAGTCTCAATTCAGGTGCGGTGAAGATGCTAAAAAGAGAA 1980  
Db 1921 TATGGCAATAGTAAACCGGAGTCTCAATTCAGGTGCGGTGAAGATGCTAAAAAGAGAA 1980  
QY 1981 GCTGACAGCTGTGAAAAAGAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA 2040  
Db 1981 GCTGACAGCTGTGAAAAAGAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA 2040  
QY 2041 CACCATGACAAATCTGTAATCTGCTGGGGCATGCACTGTGAGGGGCGAGTGTACTTTG 2100  
Db 2041 CACCATGACAAATCTGTAATCTGCTGGGGCATGCACTGTGAGGGGCGAGTGTACTTTG 2100  
QY 2101 ATTTTGAATATTTGCTGATGTTGCTCTCAACTACCTAAAGATGAAAGAGAGAG 2160  
Db 2101 ATTTTGAATATTTGCTGATGTTGCTCTCAACTACCTAAAGATGAAAGAGAGAG 2160  
QY 2161 TTTTCAAGGACATGCAAGAGATTTTAAAGGAAATATTTTCAAGTTTCTTACCTACTTTC 2220  
Db 2161 TTTTCAAGGACATGCAAGAGATTTTAAAGGAAATATTTTCAAGTTTCTTACCTACTTTC 2220  
QY 2221 CAGGCACATTTCAAAATTCAGCATGCTGGTTTCAAGAGGTTTCAAGTTTCAAGTTTTC 2280  
Db 2221 CAGGCACATTTCAAAATTCAGCATGCTGGTTTCAAGAGGTTTCAAGTTTCAAGTTTTC 2280  
QY 2281 GATCAGCTCTCAGGTTTCAATGGGAATTTCAATTTCTGAGATGAGATTAATGAA 2340  
Db 2281 GATCAGCTCTCAGGTTTCAATGGGAATTTCAATTTCTGAGATGAGATTAATGAA 2340  
QY 2341 AACAGAGAGGCTGGCAG 2400  
Db 2341 AACAGAGAGAGGCTGGCAG 2400  
QY 2401 CTCCTTTCTTTGGGTACCAAGTGGCCAAAGGATGGAATTTCTTGGAGTTCAAGTCTGT 2460  
Db 2401 CTCCTTTCTTTGGGTACCAAGTGGCCAAAGGATGGAATTTCTTGGAGTTCAAGTCTGT 2460  
QY 2461 GTCCACAGAGACTGTCAGGCAAGGAAATGTGTTGTTGTCACCCAGGAGAGGTTGAGATC 2520  
Db 2461 GTCCACAGAGACTGTCAGGCAAGGAAATGTGTTGTTGTTGTCACCCAGGAGAGGTTGAGATC 2520  
QY 2521 TGTGATTTTGGACTGGCCCGGAGACATCTGAGCGACTCCAGTACGTCTGTCAGGGGCAAC 2580  
Db 2521 TGTGATTTTGGACTGGCCCGGAGACATCTGAGCGACTCCAGTACGTCTGTCAGGGGCAAC 2580

```
QY 2581 GCACGGCTCGCGTGAAGTGGATGGCCACCGGAGCTTATTTGAAGGATCTACACAATC 2640
DB 2581 GCACGGCTCGCGTGAAGTGGATGGCCACCGGAGCTTATTTGAAGGATCTACACAATC 2640
QY 2641 AAGAGTGAGCTCTGGTCTTACGGCATCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700
DB 2641 AAGAGTGAGCTCTGGTCTTACGGCATCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700
QY 2701 CCTTACCTGGGCAATCTCTGTGAGCGTAACTTCTTATAAACTGAATCAGAGTGGATTTAA 2760
DB 2701 CCTTACCTGGGCAATCTCTGTGAGCGTAACTTCTTATAAACTGAATCAGAGTGGATTTAA 2760
QY 2761 ATGGAGCAGCAATCTTATGCCACAGAGGATATCTTTGTAATGCAATCTCTGGCT 2820
DB 2761 ATGGAGCAGCAATCTTATGCCACAGAGGATATCTTTGTAATGCAATCTCTGGCT 2820
QY 2821 TTTGACTCAAGGAAGCGGCATCTTCCCAACCTGACTTCAATTTAGGATGTGAGCTG 2880
DB 2821 TTTGACTCAAGGAAGCGGCATCTTCCCAACCTGACTTCAATTTAGGATGTGAGCTG 2880
QY 2881 GCAGAGGAGAGAGCATGTATCAAGATCCATCCATCTACCAAAACAGCGGCCCT 2940
DB 2881 GCAGAGGAGAGAGCATGTATCAAGATCCATCCATCTACCAAAACAGCGGCCCT 2940
QY 2941 CAGCAGAGAGCGGCTCAGAGCCAGTCGCGACAGCGCCAGGTGAAGTTACAGAGAA 3000
DB 2941 CAGCAGAGAGCGGCTCAGAGCCAGTCGCGACAGCGCCAGGTGAAGTTACAGAGAA 3000
QY 3001 AGAAGTTAGCAGGAGCGCTTGGACCCCGCCACCTAGCAGGCTGTAGACCCAGAGCA 3060
DB 3001 AGAAGTTAGCAGGAGCGCTTGGACCCCGCCACCTAGCAGGCTGTAGACCCAGAGCA 3060
QY 3061 AGATTAGCTCGCTCTCAGGAAGCGCTCAGCGCGTGTCTCGCTGGACTTTCTCT 3120
DB 3061 AGATTAGCTCGCTCTCAGGAAGCGCTCAGCGCGTGTCTCGCTGGACTTTCTCT 3120
QY 3121 AGATGCTCTGCGCAATCTCCAAAGTCACTTCTATAAAATCAACCTCTCTCGCAG 3180
DB 3121 AGATGCTCTGCGCAATCTCCAAAGTCACTTCTATAAAATCAACCTCTCTCGCAG 3180
QY 3181 GCGGGAGAGCAATAATGAGACTTGTGTGAGCGCGCTACCTCGGGGCTTTTCCACG 3240
DB 3181 GCGGGAGAGCAATAATGAGACTTGTGTGAGCGCGCTACCTCGGGGCTTTTCCACG 3240
QY 3241 AGCTTGAGGGGAAAGCCATGTATCTGAAATATAGTATATCTTGTAAATACGTGAACAA 3300
DB 3241 AGCTTGAGGGGAAAGCCATGTATCTGAAATATAGTATATCTTGTAAATACGTGAACAA 3300
QY 3301 ACCAAACCCGTTTTTGTCTAAGGAAAGCTAAATATGATTTTAAATACTATGTTTAA 3360
DB 3301 ACCAAACCCGTTTTTGTCTAAGGAAAGCTAAATATGATTTTAAATACTATGTTTAA 3360
QY 3361 AATACTATGTAATCTTTTCACTATTTAGTGATATATTTTATGATGGAATAAATCTTC 3420
DB 3361 AATACTATGTAATCTTTTCACTATTTAGTGATATATTTTATGATGGAATAAATCTTC 3420
QY 3421 TACTGTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3453
DB 3421 TACTGTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3453
```

```
RESULT 8
PCT-US92-02750-1
; Sequence 1, Application PC/IUS9202750
; GENERAL INFORMATION:
; APPLICANT: LEMISCHKA, IHOR R.
; TITLE OF INVENTION: Totipotent Hematopoietic Stem Cell
; TITLE OF INVENTION: Receptors And Their Ligands
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCLONE SYSTEMS INCORPORATED
; STREET: 180 VARICK STREET
```

```
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: US
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02750
; FILING DATE: 19920402
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FEIT, IRVING N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-PPPT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3453 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31...3009
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 31...3006
; PCT-US92-02750-1
```

```
Query Match 100.0%; Score 3453; DB 5; Length 3453;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3453; Conservative 0; Mismatches 0; Gaps 0;

QY 1 GCGGCTGGCTTACCGCGCTCCCGAGGCCATGGGGCGTTGGCGCAGCGAGCCGCGG 60
DB 1 GCGGCTGGCTTACCGCGCTCCCGAGGCCATGGGGCGTTGGCGCAGCGAGCCGCGG 60
QY 61 CGGCTGCTGCTGCTGTTGTTTGTGAGTAAATGATTTCTTGAGACCGTTTACAAACCAAGAC 120
DB 61 CGGCTGCTGCTGCTGTTGTTTGTGAGTAAATGATTTCTTGAGACCGTTTACAAACCAAGAC 120
QY 121 CTGCTGTGATCAAGTGTTGTTTAAATCAGTCATGAGAAACAATGCTCATCGCGGGAAG 180
DB 121 CTGCTGTGATCAAGTGTTGTTTAAATCAGTCATGAGAAACAATGCTCATCGCGGGAAG 180
QY 181 CCATCATCGTACCGAATGGTGGAGGATCCCGAGAGACCTCCAGTGTACCCGAGGCGC 240
DB 181 CCATCATCGTACCGAATGGTGGAGGATCCCGAGAGACCTCCAGTGTACCCGAGGCGC 240
QY 241 CAGAGTGAAGGAGCGGTATATGAAGCGGCCACCGTGGAGGTGGCGGAGTCTGGGTCCATC 300
DB 241 CAGAGTGAAGGAGCGGTATATGAAGCGGCCACCGTGGAGGTGGCGGAGTCTGGGTCCATC 300
QY 301 ACCCTGCAAGTGAGCTGCGCAACCCAGGGGACCTTTCTGCTCTGGTCTTTAAGCAC 360
DB 301 ACCCTGCAAGTGAGCTGCGCAACCCAGGGGACCTTTCTGCTCTGGTCTTTAAGCAC 360
QY 361 AGCTCCCTGGCTGCCAGCGCATTTGATTATACAAACAGAGGAATCGTTTCCATGCGC 420
DB 361 AGCTCCCTGGCTGCCAGCGCATTTGATTATACAAACAGAGGAATCGTTTCCATGCGC 420
QY 421 ATCTTGAACTGTGACAGAGACCCAGGAGGAGATACCTACTCCATATTTCAAGGAGCGC 480
DB 421 ATCTTGAACTGTGACAGAGACCCAGGAGGAGATACCTACTCCATATTTCAAGGAGCGC 480
QY 481 GCCAACTACACAGTCTGTTTCAAGTGAATGTAAAGAGATACAGAGTGTATGTCTAAGG 540
DB 481 GCCAACTACACAGTCTGTTTCAAGTGAATGTAAAGAGATACAGAGTGTATGTCTAAGG 540
```

Db 481 GCCAATACACAGTACTGTTTACAGTGNATGTAAGAGATACACAGCTGTATGTCTAAGG 540  
Qy 541 AGACCTTTACTTTAGGAAGATGGAACCCAGAGTGCATGCTCTGATCTCCGAGGGTGT 600  
Db 541 AGACCTTTACTTTAGGAAGATGGAACCCAGAGTGCATGCTCTGATCTCCGAGGGTGT 600  
Qy 601 CCGGAGCCCACTGGGAGTGGTCTCTGACGTCCACAGGGAAGCTGTAAGAGAA 660  
Db 601 CCGGAGCCCACTGGGAGTGGTCTCTGACGTCCACAGGGAAGCTGTAAGAGAA 660  
Qy 661 GGCCCTGCTGTGTGACAGAGAGGAAAGCTACTTTCATGAGTTGTTCCGGAACGACATC 720  
Db 661 GGCCCTGCTGTGTGACAGAGAGGAAAGCTACTTTCATGAGTTGTTCCGGAACGACATC 720  
Qy 721 AGATGCTGTGTAGAAATGCACTGGGCGCGAATGCAACCAAGCTGTTCACCATGATCTA 780  
Db 721 AGATGCTGTGTAGAAATGCACTGGGCGCGAATGCAACCAAGCTGTTCACCATGATCTA 780  
Qy 781 AACAGGCTCTCAGAGCACACTGCCCGAGTTATTCCTGAAAGTGGGGAAACCCCTTGTGG 840  
Db 781 AACAGGCTCTCAGAGCACACTGCCCGAGTTATTCCTGAAAGTGGGGAAACCCCTTGTGG 840  
Qy 841 ATCAGGTGTAAGGCACTCATGTGAAACCATGATTCGGGCTCACCTGGGAGCTGGAAGAC 900  
Db 841 ATCAGGTGTAAGGCACTCATGTGAAACCATGATTCGGGCTCACCTGGGAGCTGGAAGAC 900  
Qy 901 AAAGCCCTGGAGGAGGCGAGTACTTTGAGATGAGTACTTACTCCACAAAAGGACCATG 960  
Db 901 AAAGCCCTGGAGGAGGCGAGTACTTTGAGATGAGTACTTACTCCACAAAAGGACCATG 960  
Qy 961 ATTCCGATTTCTTTGGCTTTGTGTCTTCCTGGGAGGAAAGCAACCGGATTTATACACC 1020  
Db 961 ATTCCGATTTCTTTGGCTTTGTGTCTTCCTGGGAGGAAAGCAACCGGATTTATACACC 1020  
Qy 1021 TGCTCTTCTCAAGACACCCAGCCAGTCAGGTTGGTGACCTACCTAGAAAAGGGTTT 1080  
Db 1021 TGCTCTTCTCAAGACACCCAGCCAGTCAGGTTGGTGACCTACCTAGAAAAGGGTTT 1080  
Qy 1081 ATAAACGCTACAGCTCGCAAGAGAGTATGAAATGTACCCGCTACGAAAAGTTCTGCTTC 1140  
Db 1081 ATAAACGCTACAGCTCGCAAGAGAGTATGAAATGTACCCGCTACGAAAAGTTCTGCTTC 1140  
Qy 1141 TCAGTCAAGTTTAAAGGCTACCCAGAAATCCGATGCGAGTGGATCTTCTCTCAAGCCTCA 1200  
Db 1141 TCAGTCAAGTTTAAAGGCTACCCAGAAATCCGATGCGAGTGGATCTTCTCTCAAGCCTCA 1200  
Qy 1201 TTTCTTGTGACAGAGAGGCTCGAGGATGGGTACAGCATATCTAATTTTGGATCAT 1260  
Db 1201 TTTCTTGTGACAGAGAGGCTCGAGGATGGGTACAGCATATCTAATTTTGGATCAT 1260  
Qy 1261 AAGAAACAGCCAGGAGAGTACATATCTATGCAAGAAATGATGAGCCCGAGTTACACAAA 1320  
Db 1261 AAGAAACAGCCAGGAGAGTACATATCTATGCAAGAAATGATGAGCCCGAGTTACACAAA 1320  
Qy 1321 ATGTTCAAGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAATGCTCTAGCCAGCCAG 1380  
Db 1321 ATGTTCAAGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAATGCTCTAGCCAGCCAG 1380  
Qy 1381 GCGTCTTCTTCTGTATGGCTACCCGCTACCCCTTGGACCTGGAAGAGTGTTCGGAC 1440  
Db 1381 GCGTCTTCTTCTGTATGGCTACCCGCTACCCCTTGGACCTGGAAGAGTGTTCGGAC 1440  
Qy 1441 AAATCTCCCAATTGACAGGAGAAATCCAGAGGAGTTTGGAAATAAGAAAGCTTAACAGA 1500  
Db 1441 AAATCTCCCAATTGACAGGAGAAATCCAGAGGAGTTTGGAAATAAGAAAGCTTAACAGA 1500  
Qy 1501 AAAGTGTGTGGCCAGTGGGTGTCGAGCAGTACTCTAAATATGAGTGGAGCCGGGAAAGG 1560  
Db 1501 AAAGTGTGTGGCCAGTGGGTGTCGAGCAGTACTCTAAATATGAGTGGAGCCGGGAAAGG 1560  
Qy 1561 CTTCTGGTCAAAATGCTGTGCTACAAATTTCTATGGCACGCTCTTGGCAAAACCATCTTTT 1620  
Db 1561 CTTCTGGTCAAAATGCTGTGCTACAAATTTCTATGGCACGCTCTTGGCAAAACCATCTTTT 1620

Qy 1621 AACTCACAGGCGCCCTTCCCTTTTCATCCAGACAACTCTCTCTTATGCGACCAATTGGG 1680  
Db 1621 AACTCACAGGCGCCCTTCCCTTTTCATCCAGACAACTCTCTCTTATGCGACCAATTGGG 1680  
Qy 1681 CTCTGTCTCCCTTTCATTTGTTGTTCTCATTTGTTGTTGATCTGCCCAAAATACAAAAGCAA 1740  
Db 1681 CTCTGTCTCCCTTTCATTTGTTGTTCTCATTTGTTGTTGATCTGCCCAAAATACAAAAGCAA 1740  
Qy 1741 TTTAGGTACGAGAGTCACTGCGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 1800  
Db 1741 TTTAGGTACGAGAGTCACTGCGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 1800  
Qy 1801 TTTACGTTGCTTTCAGGAGTATGAAATATGAAATATGAAATATGAAATATGAAATATGAAATAT 1860  
Db 1801 TTTACGTTGCTTTCAGGAGTATGAAATATGAAATATGAAATATGAAATATGAAATATGAAATAT 1860  
Qy 1861 TTAGAGTTTGGGAAAGGCTCTGGGCTTTCGGGAGGCTTTCGGGAGGCTTTCGGGAGGCTTTCGGG 1920  
Db 1861 TTAGAGTTTGGGAAAGGCTCTGGGCTTTCGGGAGGCTTTCGGGAGGCTTTCGGGAGGCTTTCGGG 1920  
Qy 1921 TATGGCAATTAGTAAACCGGAGTCTCAATTCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 1980  
Db 1921 TATGGCAATTAGTAAACCGGAGTCTCAATTCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 1980  
Qy 1981 GCTGACAGCTGTGAAAGAAAGCTCTCATGTCGAGCTCAAAATGATGACCCACCTGGGA 2040  
Db 1981 GCTGACAGCTGTGAAAGAAAGCTCTCATGTCGAGCTCAAAATGATGACCCACCTGGGA 2040  
Qy 2041 CACCATGACAACATCGTGAATCTGCTGGGGCATGCACTGTCAGGCGAGTGTACTTTG 2100  
Db 2041 CACCATGACAACATCGTGAATCTGCTGGGGCATGCACTGTCAGGCGAGTGTACTTTG 2100  
Qy 2101 ATTTTGAATATTTGCTGTGCTGACCTCTCAATTAAGGAAATATTTTCACTTCTTCTTCTTCTTCT 2160  
Db 2101 ATTTTGAATATTTGCTGTGCTGACCTCTCAATTAAGGAAATATTTTCACTTCTTCTTCTTCTTCT 2160  
Qy 2161 TTTTCAAGGACATGGAAGAGATTTTAAAGGAAATATTTTCACTTCTTCTTCTTCTTCTTCTTCTTCT 2220  
Db 2161 TTTTCAAGGACATGGAAGAGATTTTAAAGGAAATATTTTCACTTCTTCTTCTTCTTCTTCTTCTTCT 2220  
Qy 2221 CAGGCACTTCAAAATTCAGCATGCTGTTTCAAGGAAATATTTTCACTTCTTCTTCTTCTTCTTCTTCT 2280  
Db 2221 CAGGCACTTCAAAATTCAGCATGCTGTTTCAAGGAAATATTTTCACTTCTTCTTCTTCTTCTTCTTCT 2280  
Qy 2281 GATCAGCTCTCAGGCTTCAATGGAATTCATTTCACTTCTGAGATGAGATTTGAAATATGAA 2340  
Db 2281 GATCAGCTCTCAGGCTTCAATGGAATTCATTTCACTTCTGAGATGAGATTTGAAATATGAA 2340  
Qy 2341 AACCAAGAGGCTGGCAAGAGAGAGAGAGAGATTTTGAACGCTGCTGACGTTTGAAGAC 2400  
Db 2341 AACCAAGAGGCTGGCAAGAGAGAGAGAGAGATTTTGAACGCTGCTGACGTTTGAAGAC 2400  
Qy 2401 CTCCTTTGCTTTGCTGCTACCAAGTGGCCAAAGGCAATGAAATTCCTGAGTTCAGTTCGTT 2460  
Db 2401 CTCCTTTGCTTTGCTGCTACCAAGTGGCCAAAGGCAATGAAATTCCTGAGTTCAGTTCGTT 2460  
Qy 2461 GTCCACAGAGACCTGGCAGCAGGAAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2520  
Db 2461 GTCCACAGAGACCTGGCAGCAGGAAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2520  
Qy 2521 TGTGACTTTGAGCTGGCCGAGACATCTCTGAGCGCATCTCCAGCTACGTCGTGAGGCGCAAC 2580  
Db 2521 TGTGACTTTGAGCTGGCCGAGACATCTCTGAGCGCATCTCCAGCTACGTCGTGAGGCGCAAC 2580  
Qy 2581 GCACGCTGCGGCTGAGTGGATGGCACCAGAGCTTTTGAAGGATCTACACAATC 2640  
Db 2581 GCACGCTGCGGCTGAGTGGATGGCACCAGAGCTTTTGAAGGATCTACACAATC 2640  
Qy 2641 AAGAGTGCAGCTGCTGCTCCTCAGGATCTCTCTGGGAGATATTTTCACTGGTGTGAC 2700  
Db 2641 AAGAGTGCAGCTGCTGCTCCTCAGGATCTCTCTGGGAGATATTTTCACTGGTGTGAC 2700





QY	661	GGCCTGCTGTTGTCAGAAAAGGAGAAAGGTACTTCA	TGAGTTGTTGCGAA	CAGACATC	720
Db	661				720
QY	721	AGATGCTGCTAGAAATGCACTGGGCGCGAATGCA	CAAGCTGTTTCA	CCATAGATCTA	780
Db	721				780
QY	781	AACAGAGCTCTCAGACGACACTGCCCCAGTTATCT	GAAAGCTGGGGAA	CCCTTTGCG	840
Db	781				840
QY	841	ATCAGGTGTAAGGCCATCATGTGAA	CCATGGAATTCGGGCTCACCTGGGAGCTGGAAGAC	900	
Db	841				900
QY	901	AAAGCCCTGGAGGAGGCGAGTACTTTGAGATGAGTAC	TCTTCTCCAAACAGGACCATG	960	
Db	901				960
QY	961	ATTCCGGATCTCTTGGGCTTTGTCCTTCGTTGGAG	AGAAACGACACCGGATATTACAC	1020	
Db	961				1020
QY	1021	TGCTCTTCTCAAAGCACCCAGCCAGTCAGCGTGTG	TGACCATCTCAGAAAAAGGGTTT	1080	
Db	1021				1080
QY	1081	ATAAACGCTACAGCTCGAAGAGAGATGAAATGCA	CCCGTACGAAAAAGTTCTGCTTC	1140	
Db	1081				1140
QY	1141	TCAGTCAGGTTTAAAGCGTACCCACGAATCGATG	CAGTGGATCTTCTCAGAGCTCA	1200	
Db	1141				1200
QY	1201	TTTCTCTGTGAAACAGAGAGGCTCGAGAGTGGT	ACAGCATATCTAAATTTTTCGATCAT	1260	
Db	1201				1260
QY	1261	AAGAACAGCCAGGAGAGTACATATTTATG	CAGAAATGATGCGCCAGTTTCA	1320	
Db	1261				1320
QY	1321	ATGTTACGCTGAAATATAGAAAGAAACCTCA	AGTGTAGCAATGCTCAGCCAGCCAG	1380	
Db	1321				1380
QY	1381	CGCTCTGTTCTCTGATGGCTACCCGCTACCTCT	TGGACCTTGGAAAGAGTGTTCGGAC	1440	
Db	1381				1440
QY	1441	AAATCTCCCAATTGCA	CGGAGGAAATCCAGAGGATTTTGGAA	TAAAAAGGCTTAACAGA	1500
Db	1441				1500
QY	1501	AAAGTTTGGCCAGTGGGTGTCGAGCAGTACTCT	AAATATGAGTGAAGCGCGGAAGGG	1560	
Db	1501				1560
QY	1561	CTTCTGGTCAAAATGCTGTCGTAATTCAT	TGCGCAGTCTTCGAAACCATCTTTT	1620	
Db	1561				1620
QY	1621	AACTCACAGGCCCTTCCCTTTT	CATCCAGACAACTCTCCTTCTATGCGACCATGGG	1680	
Db	1621				1680
QY	1681	CTCTGCTCCCTTCATTTGTTCTCATATGTTGT	GTCTGCCCAAAATACAAAAGCAA	1740	
Db	1681				1740

Qy	1741	TTTTAGTACAGAGT	CAGCTG	CAGATG	ATCCAGGTG	ACTGGCCCTCGATACGAGTAC	1801	
Db	1741	TTTTAGTACAGAGT	CAGCTG	CAGATG	ATCCAGGTG	ACTGGCCCTCGATACGAGTAC	1800	
Qy	1801	TTCTACGTTTCAC	TTTCAGGG	ACTATGA	ATTAATG	ATTAATGATGAACTTAAAGTGGAGTTCCCGAGAGAA	1860	
Db	1801	TTCTACGTTTCAC	TTTCAGGG	ACTATGA	ATTAATG	ATTAATGATGAACTTAAAGTGGAGTTCCCGAGAGAA	1860	
Qy	1861	TTAGAGTTTGGG	AAGGTCCT	GGGGTCTGG	CGCTTTTCGGGAGGGTGA	TGAACGCAACGGCC	1920	
Db	1861	TTAGAGTTTGGG	AAGGTCCT	GGGGTCTGG	CGCTTTTCGGGAGGGTGA	TGAACGCAACGGCC	1920	
Qy	1921	TATGGCAATAGT	ATAAACCGG	AGTCTCAA	TTTCAGGTGGCGGTGA	TGCTCTAAAGAGAAA	1980	
Db	1921	TATGGCAATAGT	ATAAACCGG	AGTCTCAA	TTTCAGGTGGCGGTGA	TGCTCTAAAGAGAAA	1980	
Qy	1981	GCTGACAGCTG	TGAAAAGAA	AGCTCTCAT	CTGCGGAGCTCA	AAATGATGACCCACCTGGGA	2040	
Db	1981	GCTGACAGCTG	TGAAAAGAA	AGCTCTCAT	CTGCGGAGCTCA	AAATGATGACCCACCTGGGA	2040	
Qy	2041	CACCATGACAA	CATCTGTA	ATCTGCTGGGGCA	TGCACATGTCAGGGCCAGTGTA	CTTG	2100	
Db	2041	CACCATGACAA	CATCGTGA	ATCTGCTGGGGCA	TGCACATGTCAGGGCCAGTGTA	CTTG	2100	
Qy	2101	ATTTTTGAAT	TTGTTGCTAT	TGCTGACCTCCT	CMACTACCTPAAGAGTAA	AAGAGAG	2160	
Db	2101	ATTTTTGAAT	TTGTTGCTAT	TGCTGACCTCCT	CMACTACCTPAAGAGTAA	AAGAGAG	2160	
Qy	2161	TTTTACAGGAC	ATGGAAGAGAT	TTTTTAAAGMA	CATAATTTTCAGTCTCTAC	CCCTACCTTTC	2220	
Db	2161	TTTTACAGGAC	ATGGAAGAGAT	TTTTTAAAGMA	CATAATTTTCAGTCTCTAC	CCCTACCTTTC	2220	
Qy	2221	CAGGCACAT	TTCAAATTCAG	CATCCCTGGT	TTTCAGGAAAT	TTTCAGTCTCTACCCCTACCTTTC	2280	
Db	2221	CAGGCACAT	TTCAAATTCAG	CATCCCTGGT	TTTCAGGAAAT	TTTCAGTCTCTACCCCTACCTTTC	2280	
Qy	2281	GATCAGCTCT	CAGGGTCAA	TTGGAA	TTCAA	TTCAATTCGAGATGAGATGGAATATGAA	2340	
Db	2281	GATCAGCTCT	CAGGGTCAA	TTGGAA	TTCAA	TTCAATTCGAGATGAGATGGAATATGAA	2340	
Qy	2341	AACAGAAGAG	CTGGCAGA	AGAGGAGG	AGATTTTGAACGTG	CTGACGTTTGAAGAC	2400	
Db	2341	AACAGAAGAG	CTGGCAGA	AGAGGAGG	AGATTTTGAACGTG	CTGACGTTTGAAGAC	2400	
Qy	2401	CTCCTTTGCT	TTTGGTACCA	AGTGGCCAA	AGGCAATGAA	TTTTCCTGGAAGTTCAAAGTCTG	2460	
Db	2401	CTCCTTTGCT	TTTGGTACCA	AGTGGCCAA	AGGCAATGAA	TTTTCCTGGAAGTTCAAAGTCTG	2460	
Qy	2461	GTCCACAGAG	ACCTGGCAG	CGAGAA	TGTTGGTCA	CCCCACGGGAAGTGGTGAAGATC	2520	
Db	2461	GTCCACAGAG	ACCTGGCAG	CGAGAA	TGTTGGTCA	CCCCACGGGAAGTGGTGAAGATC	2520	
Qy	2521	TGTGACTTTG	AGCTGGCCG	GAGACAT	CTCTGAGCGA	CTCCAGCTACGTCGTCAGGGGCAAC	2580	
Db	2521	TGTGACTTTG	AGCTGGCCG	GAGACAT	CTCTGAGCGA	CTCCAGCTACGTCGTCAGGGGCAAC	2580	
Qy	2581	GCACGGCTGC	CGGTGAAGT	GTGACCG	AGGCAATGAA	TTTGAAGGATCTACA	CAATC	2640
Db	2581	GCACGGCTGC	CGGTGAAGT	GTGACCG	AGGCAATGAA	TTTGAAGGATCTACA	CAATC	2640
Qy	2641	AAGAGTGAC	GTCTGGT	CTCAGG	CATCTCTCTGGGAGAT	TTTTTCACTGGGGTGAAC	2700	
Db	2641	AAGAGTGAC	GTCTGGT	CTCAGG	CATCTCTCTGGGAGAT	TTTTTCACTGGGGTGAAC	2700	
Qy	2701	CCTTACCTTG	GCAATTCCT	GTGCGAG	CAATCTTCTAT	ATAA	CTGATTCAGAGTGGATTTTAA	2760
Db	2701	CCTTACCTTG	GCAATTCCT	GTGCGAG	CAATCTTCTAT	ATAA	CTGATTCAGAGTGGATTTTAA	2760
Qy	2761	ATGGAGAGC	CAATCTAT	GCCACAGAG	GGATATAC	TTTGAATGCAATCTCTGCTGGCT	2820	
Db	2761	ATGGAGAGC	CAATCTAT	GCCACAGAG	GGATATAC	TTTGAATGCAATCTCTGCTGGCT	2820	
Qy	2821	TTTGACTCA	AGAGCGGCC	CACTCTT	CCCCAACCTG	CACTTTTTCAGGATGTCAGCTG	2880	

Db 2821 TTTGACTCAAGAGCGGCATCTTCCCAACTGACTTCAATTTTAGATGTCAGCTG 2880  
QY 2881 GCAGAGGAGAGAGCATGTATCAACAATCCATCCATCAACCAAGAGGCGCCCT 2940  
Db 2881 GCAGAGGAGAGAGCATGTATCAACAATCCATCCATCAACCAAGAGGCGCCCT 2940  
QY 2941 CAGCAGAGAGGCGGCTCAGAGCCAGTCCGACAGCGCCAGGTGAAGATTACAGAGAA 3000  
Db 2941 CAGCAGAGAGGCGGCTCAGAGCCAGTCCGACAGCGCCAGGTGAAGATTACAGAGAA 3000  
QY 3001 AGAAGTTAGCAGAGGCGCTTGGACCCCGCCAGTCCGACAGCGCTGAGACCGCAGACCA 3060  
Db 3001 AGAAGTTAGCAGAGGCGCTTGGACCCCGCCAGTCCGACAGCGCTGAGACCGCAGACCA 3060  
QY 3061 AGATTAGCTCGCCCTCTGAGGAGCGCCCTACAGCGCGTGTCTCGCTGGACTTTCTCT 3120  
Db 3061 AGATTAGCTCGCCCTCTGAGGAGCGCCCTACAGCGCGTGTCTCGCTGGACTTTCTCT 3120  
QY 3121 AGATGCTCTGCCATTACTCCAAAGTGAATCTTATAAAATCAAACTCTCTCGCCAG 3180  
Db 3121 AGATGCTCTGCCATTACTCCAAAGTGAATCTTATAAAATCAAACTCTCTCGCCAG 3180  
QY 3181 GCGGAGAGCCAAATAGAGACTTGTGGTGAAGCGGCTACCTGGGGGCTTTCCACG 3240  
Db 3181 GCGGAGAGCCAAATAGAGACTTGTGGTGAAGCGGCTACCTGGGGGCTTTCCACG 3240  
QY 3241 AGCTTGAGGAGAAAGCCATGATCTGAATATAGTATATCTTGTAAATAGTGAAACAA 3300  
Db 3241 AGCTTGAGGAGAAAGCCATGATCTGAATATAGTATATCTTGTAAATAGTGAAACAA 3300  
QY 3301 ACCAAACCCGTTTTTGTAGGAGAAAGCTAAATATGATTTTTTAAATATCTATGTTTAA 3360  
Db 3301 ACCAAACCCGTTTTTGTAGGAGAAAGCTAAATATGATTTTTTAAATATCTATGTTTAA 3360  
QY 3361 AATCTATGTAACCTTTTCTATCTATTTAGTATATATTTATGATGGAATAAACTTTC 3420  
Db 3361 AATCTATGTAACCTTTTCTATCTATTTAGTATATATTTATGATGGAATAAACTTTC 3420  
QY 3421 TACTGTAAATAAAAAAAAAAAAAAAAAAAAAA 3453  
Db 3421 TACTGTAAATAAAAAAAAAAAAAAAAAAAAAA 3453

RESULT 10  
PCT-US92-09893-1  
Sequence 1, Application PC/TUS9209893  
GENERAL INFORMATION:  
APPLICANT: Lemischka, Ihor R.  
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ImClone Systems Incorporated  
STREET: 180 Varick Street  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10014  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/09893  
FILING DATE: 19921116  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Feit, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-7PT  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3453 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
FRAGMENT TYPE: N-terminal  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 112..3006  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 31..111  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 31..3009  
PCT-US92-09893-1  
Query Match 100.0%; Score 3453; DB 5; Length 3453;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCGGCTTGGCTTACCGCGCTCCGAGGCGCATCGGGCGTTGGCGCAGGCGACCGG 60  
Db 1 GCGGCTTGGCTTACCGCGCTCCGAGGCGCATCGGGCGTTGGCGCAGGCGACCGG 60  
QY 61 GCGGCTTGGCTTGGTTTGTCTAGTAAATGATTTCTGAGACCGTTTCAAAACCAAG 120  
Db 61 GCGGCTTGGCTTGGTTTGTCTAGTAAATGATTTCTGAGACCGTTTCAAAACCAAG 120  
QY 121 CTGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAACAAATGCTCATCGCGGAAAG 180  
Db 121 CTGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAACAAATGCTCATCGCGGAAAG 180  
QY 181 CCATCATCGTACCGAATGGTGGGAGGATCCCGAGAACACCTCCAGTGTACCCGAGGCG 240  
Db 181 CCATCATCGTACCGAATGGTGGGAGGATCCCGAGAACACCTCCAGTGTACCCGAGGCG 240  
QY 241 CAGAGTGAAGGAGCGGTATATGAAGCGGCCACCGTGGAGTGGCGGAGTCTGGTCCATC 300  
Db 241 CAGAGTGAAGGAGCGGTATATGAAGCGGCCACCGTGGAGTGGCGGAGTCTGGTCCATC 300  
QY 301 ACCCTGAAAGTGCAGTCTGCGCACCCAGGCGACCTTTCTGCTCTGGGTCTTTAAGCAC 360  
Db 301 ACCCTGAAAGTGCAGTCTGCGCACCCAGGCGACCTTTCTGCTCTGGGTCTTTAAGCAC 360  
QY 361 AGCTCCCTGGGCTGCCAGCGGCACTTTGATTTTCAAAACAGAGAAATCGTTTCCATGCC 420  
Db 361 AGCTCCCTGGGCTGCCAGCGGCACTTTGATTTTCAAAACAGAGAAATCGTTTCCATGCC 420  
QY 421 ATCTTGAACGTGACAGAGACCCAGGCGAGGATACCTACTCCATATTCAGAGCGAACGC 480  
Db 421 ATCTTGAACGTGACAGAGACCCAGGCGAGGATACCTACTCCATATTCAGAGCGAACGC 480  
QY 481 GCCAACTACACAGTACTGTTTCAAGTGAATGAAGAGATACACAGCTGTATGTCTAAGG 540  
Db 481 GCCAACTACACAGTACTGTTTCAAGTGAATGAAGAGATACACAGCTGTATGTCTAAGG 540  
QY 541 AGACCTTACTTTAGGAGATGAAAAACAGAGTGAACCTGCTCTGCTCTCCAGGCGTGT 600  
Db 541 AGACCTTACTTTAGGAGATGAAAAACAGAGTGAACCTGCTCTGCTCTCCAGGCGTGT 600  
QY 601 CCGGAGCCCACTGTGGAGTGGTCTCTGACCTCCACAGGAAAGCTGTAAAGAA 660  
Db 601 CCGGAGCCCACTGTGGAGTGGTCTCTGACCTCCACAGGAAAGCTGTAAAGAA 660  
QY 661 GGCCTCTGTGTTCAGAAAGGAGAAAGGTACTTTCATGAGTGTTCGGAACAGACATC 720  
Db 661 GGCCTCTGTGTTCAGAAAGGAGAAAGGTACTTTCATGAGTGTTCGGAACAGACATC 720

Db 661 GGCCCTGCTGTGTGTCAGAAAGGAGAAAGGTACTTTCATGAGTTGTTTCGGAAACAGACATC 720  
Qy 721 AGATGCTGTCTAGATGACTGGCGCGGAATGACCAAGCTGTTCACCATAGATCTA 780  
Db 721 AGATGCTGTCTAGAAATGACTGGCGCGGAATGACCAAGCTGTTCACCATAGATCTA 780  
Qy 781 AACCAGGCTCCTCAGAGCACACTGCCCCAGTTATTCCTGAAAGTGGGGGAACCCCTTGTGG 840  
Db 781 AACCAGGCTCCTCAGAGCACACTGCCCCAGTTATTCCTGAAAGTGGGGGAACCCCTTGTGG 840  
Qy 841 ATCAGGTGAAGGCATCCATGTGAACCATGTGAATTCGGGTCACTCCCAAAACAGAGACCATG 900  
Db 841 ATCAGGTGAAGGCATCCATGTGAACCATGTGAATTCGGGTCACTCCCAAAACAGAGACCATG 900  
Qy 901 AAAGCCCTGAGGAGGGCAGTACTTTTGTGATGAGTACTTACTCCCAAAACAGAGACCATG 960  
Db 901 AAAGCCCTGAGGAGGGCAGTACTTTTGTGATGAGTACTTACTCCCAAAACAGAGACCATG 960  
Qy 961 ATTCCGATTCCTTTGGCTTTTGTCTTCCTGGGAAAGGAAACGACACCGGATATTACACC 1020  
Db 961 ATTCCGATTCCTTTGGCTTTTGTCTTCCTGGGAAAGGAAACGACACCGGATATTACACC 1020  
Qy 1021 TGCTCTTCTCAAGCACCCAGCCAGTCAAGCGTTGGTGAACCATCTTAGAAAAGGGTTT 1080  
Db 1021 TGCTCTTCTCAAGCACCCAGCCAGTCAAGCGTTGGTGAACCATCTTAGAAAAGGGTTT 1080  
Qy 1081 ATAAACGCTACCGTCCGAAAGAGTATGAATTCGACCGTGAACCAAGTCTGCTTC 1140  
Db 1081 ATAAACGCTACCGTCCGAAAGAGTATGAATTCGACCGTGAACCAAGTCTGCTTC 1140  
Qy 1141 TCAGTCAGGTTTAAAGGTTACCCAGCAATCCGATGACGTTGATCTTCTCAAGCCCTCA 1200  
Db 1141 TCAGTCAGGTTTAAAGGTTACCCAGCAATCCGATGACGTTGATCTTCTCAAGCCCTCA 1200  
Qy 1201 TTTCTCTGTGAACAGAGAGGCTGGAGGATGGGTACAGCATATCTAAATTTTGGATCAT 1260  
Db 1201 TTTCTCTGTGAACAGAGAGGCTGGAGGATGGGTACAGCATATCTAAATTTTGGATCAT 1260  
Qy 1261 AAGAACAGCCAGAGAGTACATATTCATGCAAAAATGATGACGCCAGTTTCAACAAA 1320  
Db 1261 AAGAACAGCCAGAGAGTACATATTCATGCAAAAATGATGACGCCAGTTTCAACAAA 1320  
Qy 1321 ATGTTCAAGCTGAATATAAGAAAGAACCTCAAGTGTAGCAAAATGCTCAGCCAGCCAG 1380  
Db 1321 ATGTTCAAGCTGAATATAAGAAAGAACCTCAAGTGTAGCAAAATGCTCAGCCAGCCAG 1380  
Qy 1381 GCGTCTGTTCTCTGATGGCTACCCGCTACCCCTCTTGGACCTGGAAAGAGTGTTCGGAC 1440  
Db 1381 GCGTCTGTTCTCTGATGGCTACCCGCTACCCCTCTTGGACCTGGAAAGAGTGTTCGGAC 1440  
Qy 1441 AAATCTCCCAATTCGACGGAGGAAATCCCAAGAGGAGTTTGGAAATAAAAGGCTAACAGA 1500  
Db 1441 AAATCTCCCAATTCGACGGAGGAAATCCCAAGAGGAGTTTGGAAATAAAAGGCTAACAGA 1500  
Qy 1501 AAAGTGTGTGGCAGTGGGTGTCAGAGAGTACTCTAAATATGATGAGGCGGGAAAGG 1560  
Db 1501 AAAGTGTGTGGCAGTGGGTGTCAGAGAGTACTCTAAATATGATGAGGCGGGAAAGG 1560  
Qy 1561 CTTCTGTCAAAATGCTGTGCTCAATTTCTATGGCAGCTCTTGGCAAAACCATCTTTTAA 1620  
Db 1561 CTTCTGTCAAAATGCTGTGCTCAATTTCTATGGCAGCTCTTGGCAAAACCATCTTTTAA 1620  
Qy 1621 AACTCAGCAGGCCCCCTTCCCTTTTATCCAAAGAACATCTCCTTTATGCGACCATTTGGG 1680  
Db 1621 AACTCAGCAGGCCCCCTTCCCTTTTATCCAAAGAACATCTCCTTTATGCGACCATTTGGG 1680  
Qy 1681 CTCGTCTCCCTTCATTTGTTTCTCATTTGTGATCTGCGCAAAATACAAAAGCAA 1740  
Db 1681 CTCGTCTCCCTTCATTTGTTTCTCATTTGTGATCTGCGCAAAATACAAAAGCAA 1740  
Qy 1741 TTTAGGTACGAGAGTCAGCTGAGATGATCCAGTGTGCTGCCCCCTGATACAGGTAC 1800  
Db 1741 TTTAGGTACGAGAGTCAGCTGAGATGATCCAGTGTGCTGCCCCCTGATACAGGTAC 1800

Qy 1801 TTCTACGTTGACTTCAGGGACTATGAAATGACCTTAAAGTGGAGTTCCCGAGAGAGAAC 1860  
Db 1801 TTCTACGTTGACTTCAGGGACTATGAAATGACCTTAAAGTGGAGTTCCCGAGAGAGAAC 1860  
Qy 1861 TTAGAGTTTGGGAAGTCTCGGGGTCTGGCGCTTTTCGGGAGGGTGTAAACCCACGGCC 1920  
Db 1861 TTAGAGTTTGGGAAGTCTCGGGGTCTGGCGCTTTTCGGGAGGGTGTAAACCCACGGCC 1920  
Qy 1921 TATGGCATTTAGTAAACCGGAGTCTCAATTCAGGTGGCGGTGAAGTGTAAAGAGAGAA 1980  
Db 1921 TATGGCATTTAGTAAACCGGAGTCTCAATTCAGGTGGCGGTGAAGTGTAAAGAGAGAA 1980  
Qy 1981 GCTGACAGCTGTCAAAAAGAGCTCTCATGTCGGAGCTCAAAATGATGACCCACTGGGA 2040  
Db 1981 GCTGACAGCTGTCAAAAAGAGCTCTCATGTCGGAGCTCAAAATGATGACCCACTGGGA 2040  
Qy 2041 CACCATGACAAACATCGTGAATCTGCTGGGGCATGCACACTGTCAGGGCCAGTGTACTTG 2100  
Db 2041 CACCATGACAAACATCGTGAATCTGCTGGGGCATGCACACTGTCAGGGCCAGTGTACTTG 2100  
Qy 2101 ATTTTGAATATTGTTGCTATGTTGACCTCTCAACTACCTAAAGATAAAGAGAGAG 2160  
Db 2101 ATTTTGAATATTGTTGCTATGTTGACCTCTCAACTACCTAAAGATAAAGAGAGAG 2160  
Qy 2161 TTTTACAGGACATGGAAGAGATTTTAAAGGAACATAATTCAGTTCTTACCCTACTTTC 2220  
Db 2161 TTTTACAGGACATGGAAGAGATTTTAAAGGAACATAATTCAGTTCTTACCCTACTTTC 2220  
Qy 2221 CAGGCACTTTCAAAATTCAGCATGCTGTTTACGAGAAGTTCAGTTTACACCCGCCCTTG 2280  
Db 2221 CAGGCACTTTCAAAATTCAGCATGCTGTTTACGAGAAGTTCAGTTTACACCCGCCCTTG 2280  
Qy 2281 GATCAGCTCTCAGGTTTCAATGGGAATTCATTTCAATCTGAAGATGAGATTGAATATGAA 2340  
Db 2281 GATCAGCTCTCAGGTTTCAATGGGAATTCATTTCAATCTGAAGATGAGATTGAATATGAA 2340  
Qy 2341 AACCAAGAGGCTGGCAGAAAGAGAGAGAGATTTTGAACCGTGTGAAGTGAAGAC 2400  
Db 2341 AACCAAGAGGCTGGCAGAAAGAGAGAGAGATTTTGAACCGTGTGAAGTGAAGAC 2400  
Qy 2401 CTCCTTTGTTTTCGTTACCAAGTGGCCAAAGGCATGGAATTTCTGGAGTTCAAGTCTGT 2460  
Db 2401 CTCCTTTGTTTTCGTTACCAAGTGGCCAAAGGCATGGAATTTCTGGAGTTCAAGTCTGT 2460  
Qy 2461 GTCCACAGAGACCTGGCAGCCAGGAATGTGTTGGTCAACCCACGGGAAGGTGGTGAAGATC 2520  
Db 2461 GTCCACAGAGACCTGGCAGCCAGGAATGTGTTGGTCAACCCACGGGAAGGTGGTGAAGATC 2520  
Qy 2521 TGTGACTTTTGGACTGCCCCGAGACATCCTGAGCGACTCCAGCTACGTCGTGAGGGCAAC 2580  
Db 2521 TGTGACTTTTGGACTGCCCCGAGACATCCTGAGCGACTCCAGCTACGTCGTGAGGGCAAC 2580  
Qy 2581 GCACGCTCCCGTGAAGTGGATGGCACCGAGAGCTTATTGAAGGGATCTACAAATC 2640  
Db 2581 GCACGCTCCCGTGAAGTGGATGGCACCGAGAGCTTATTGAAGGGATCTACAAATC 2640  
Qy 2641 AAGAGTGAAGTGTGCTTACGCGCATCTTCTCTGGAGATATTTTCACTGGGTGTGAAC 2700  
Db 2641 AAGAGTGAAGTGTGCTTACGCGCATCTTCTCTGGAGATATTTTCACTGGGTGTGAAC 2700  
Qy 2701 CCTTACCTCTGGCATTCCTGTGACCGCTAACTTCTATAAAGTGAATTCAGAGTGAATTTAAA 2760  
Db 2701 CCTTACCTCTGGCATTCCTGTGACCGCTAACTTCTATAAAGTGAATTCAGAGTGAATTTAAA 2760  
Qy 2761 ATGAGCAGCCATTTCTATGCCAAGAAAGGATATATCTTTGTAATGCAATCTCTGCTGGGCT 2820  
Db 2761 ATGAGCAGCCATTTCTATGCCAAGAAAGGATATATCTTTGTAATGCAATCTCTGCTGGGCT 2820  
Qy 2821 TTTTACTCAGGAAGCGGCCATCTTCCCCCACTGACTTCATTTTAGGATGTGAGCTG 2880  
Db 2821 TTTTACTCAGGAAGCGGCCATCTTCCCCCACTGACTTCATTTTAGGATGTGAGCTG 2880



181 CCATCATCGTACCAATGGTGGAGGATCCCAAGAGACCTCCAGTGTACCCCGAGGCGC 240  
182 |  
232 CCATCATCGTACCAATGGTGGAGGATCCCAAGAGACCTCCAGTGTACCCCGAGGCGC 291  
233 |  
241 CAGAGTGAAGGAGCGGTATATGAAGCGGCCACCGTGGAGGTGGCCGAGTCTGGGTCCATC 300  
242 |  
292 CAGAGTGAAGGAGCGGTATATGAAGCGGCCACCGTGGAGGTGGCCGAGTCTGGGTCCATC 351  
293 |  
301 ACCCTGCAAGTGGAGTCCGACACCCAGGAGGACCTTCTGCTCTGGGTCTTTAAGCAC 360  
302 |  
352 ACCCTGCAAGTGGAGTCCGACACCCAGGAGGACCTTCTGCTCTGGGTCTTTAAGCAC 411  
353 |  
361 AGCTCCCTGGGTGGCCAGCGGCATCTTGTATTTACAAAACAGAGGAATCGTTTCCATGGCC 420  
362 |  
412 AGCTCCCTGGGTGGCCAGCGGCATCTTGTATTTACAAAACAGAGGAATCGTTTCCATGGCC 471  
413 |  
421 ATCTTGAACGTGACAGAGACCCAGGAGGAGAACTACTCTCATATTCAGAGCGAACGC 480  
422 |  
472 ATCTTGAACGTGACAGAGACCCAGGAGGAGAACTACTCTCATATTCAGAGCGAACGC 531  
473 |  
481 GCCAACTACACAGTACTGTTACAGTGAATGTAAAGATACACAGCTGTATGTCTAAGG 540  
482 |  
532 GCCAACTACACAGTACTGTTACAGTGAATGTAAAGATACACAGCTGTATGTCTAAGG 591  
533 |  
541 AGACCTTACTTTAGGAAGATGGAACACCAAGATGCACTGCTCTGCACTCTCCGAGGCGT 600  
542 |  
592 AGACCTTACTTTAGGAAGATGGAACACCAAGATGCACTGCTCTGCACTCTCCGAGGCGT 651  
593 |  
601 CGGAGGCCACTGTGAGTGGGTGCTCTGCACTCCCAAGGAGAACTGTAAAGAGAA 660  
602 |  
652 CGGAGGCCACTGTGAGTGGGTGCTCTGCACTCCCAAGGAGAACTGTAAAGAGAA 711  
653 |  
661 GGCCCTGCTGTTGTCAAGAGGAGGAAAGGTAATTCATGAGTGTTCGGAAACAGACATC 720  
662 |  
712 GGCCCTGCTGTTGTCAAGAGGAGGAAAGGTAATTCATGAGTGTTCGGAAACAGACATC 771  
713 |  
721 AGATCTGTGCTAGAAATGCACTGGCCCGGAAAGGTAATTCATGAGTGTTCGGAAACAGACATC 780  
722 |  
772 AGATCTGTGCTAGAAATGCACTGGCCCGGAAAGGTAATTCATGAGTGTTCGGAAACAGACATC 831  
773 |  
781 AACAGGCTCTCAGAGCACCTGCCCGCAGTTATTCCTGAAAGTGGGGAAACCTTTGGG 840  
782 |  
832 AACAGGCTCTCAGAGCACCTGCCCGCAGTTATTCCTGAAAGTGGGGAAACCTTTGGG 891  
833 |  
841 ATCAGGTGAAGGCCATCCATGTGAACCAATGATGGGCTCACCTGGAGCTGGAAGAC 900  
842 |  
892 ATCAGGTGAAGGCCATCCATGTGAACCAATGATGGGCTCACCTGGAGCTGGAAGAC 951  
893 |  
901 AAAGCCCTGGAGGAGGAGCTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATG 960  
902 |  
952 AAAGCCCTGGAGGAGGAGCTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATG 1011  
953 |  
961 ATTGGATTTCTTGGCCTTTGCTCTCCGTTGGAGGAAACGACCGGATATTACACC 1020  
962 |  
1012 ATTGGATTTCTTGGCCTTTGCTCTCCGTTGGAGGAAACGACCGGATATTACACC 1071  
1013 |  
1021 TGCTCTTCTCAAAGACCCAGCAGTCAAGGCTGGTGGACCATCTAGAAAAGGGTTT 1080  
1022 |  
1072 TGCTCTTCTCAAAGACCCAGCAGTCAAGGCTGGTGGACCATCTAGAAAAGGGTTT 1131  
1073 |  
1081 ATAAACGCTACAGCTCGGAAGAGATGAAATTTGACCCGTTAGAAAAGGGTTT 1140  
1082 |  
1132 ATAAACGCTACAGCTCGGAAGAGATGAAATTTGACCCGTTAGAAAAGGGTTT 1191  
1133 |  
1141 TCAGTCAAGTTTAAAGCGTACCCAGGATCCGATCCGATCGAGTGGATCTTCTCAAGCCTCA 1200  
1142 |  
1192 TCAGTCAAGTTTAAAGCGTACCCAGGATCCGATCCGATCGAGTGGATCTTCTCAAGCCTCA 1251  
1193 |  
1201 TTTCCCTGTGAACAGAGAGGCTGGAGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260  
1202 |  
1252 TTTCCCTGTGAACAGAGAGGCTGGAGATGGGTACAGCATATCTAAATTTTGGCATCAT 1311  
1253 |  
1261 AAGAACAGCCAGGAGATACATATTTCTATGAGAAAATGATGACGCCCGCTTACCAA 1320  
1262 |

1312 AAGAACAGCCAGAGATACATATTTCTATGCAAAAATGATGACGCCAGTTTCCAAA 1371  
1313 |  
1321 ATGTTTCACTGATTAATAAGAAACAACTCAAGTGTCTAGCAAAATGCTCAGCCAGCCAG 1380  
1322 |  
1372 ATGTTTCACTGATTAATAAGAAACAACTCAAGTGTCTAGCAAAATGCTCAGCCAGCCAG 1431  
1373 |  
1381 GGCTCTGTTCTCTGATGGCTACCGCTACCTCTTGGACCTTGGAAAGATGTTTGGAC 1440  
1382 |  
1432 GGCTCTGTTCTCTGATGGCTACCGCTACCTCTTGGACCTTGGAAAGATGTTTGGAC 1491  
1433 |  
1441 AATCTCCCAATTTGCAAGGAGAAATCCCAAGAGGAGTTTGGAAATAAAAAGGCTTAAACA 1500  
1442 |  
1492 AATCTCCCAATTTGCAAGGAGAAATCCCAAGAGGAGTTTGGAAATAAAAAGGCTTAAACA 1551  
1493 |  
1501 AAGAGTGTGGCCAGTGGGTGTGAGCAGTACTCTAAATATATGAGTGAAGGCGGAAAGGG 1560  
1502 |  
1552 AAGAGTGTGGCCAGTGGGTGTGAGCAGTACTCTAAATATATGAGTGAAGGCGGAAAGGG 1611  
1553 |  
1561 CTCTGGTCAAAATGCTGTGGGTACAAATTTCTATGGGCAAGTCTTGGGAAACATCTTTTTTA 1620  
1562 |  
1612 CTCTGGTCAAAATGCTGTGGGTACAAATTTCTATGGGCAAGTCTTGGGAAACATCTTTTTTA 1671  
1613 |  
1621 AACTCACAGGCGCCCTTCCCTTTTCTATGGGCAAGTCTTGGGAAACATCTTTTTTA 1680  
1622 |  
1672 AACTCACAGGCGCCCTTCCCTTTTCTATGGGCAAGTCTTGGGAAACATCTTTTTTA 1731  
1673 |  
1681 CTCTGTCTCCCTTCAATTTGTTTCTCATTCAGAGCAACATCTCTCTTCTATGCGACCATTTGG 1740  
1682 |  
1732 CTCTGTCTCCCTTCAATTTGTTTCTCATTCAGAGCAACATCTCTCTTCTATGCGACCATTTGG 1791  
1733 |  
1741 TTTAGGTACAGAGTCACTGAGATGATCCAGGTGACTGGGCGGCTTGGATAACGAGTAC 1800  
1742 |  
1792 TTTAGGTACAGAGTCACTGAGATGATCCAGGTGACTGGGCGGCTTGGATAACGAGTAC 1851  
1793 |  
1801 TTTAGGTGACTTCAAGGACTATGAAATGACTTTAAGTGGAGTTCGCGAGAGAGAAC 1860  
1802 |  
1852 TTTAGGTGACTTCAAGGACTATGAAATGACTTTAAGTGGAGTTCGCGAGAGAGAAC 1911  
1853 |  
1861 TTTAGGTGACTTCAAGGACTATGAAATGACTTTAAGTGGAGTTCGCGAGAGAGAAC 1920  
1862 |  
1912 TTTAGGTGACTTCAAGGACTATGAAATGACTTTAAGTGGAGTTCGCGAGAGAGAAC 1971  
1913 |  
1921 TTTAGGTGACTTCAAGGACTATGAAATGACTTTAAGTGGAGTTCGCGAGAGAGAAC 1980  
1922 |  
1972 TTTAGGTGACTTCAAGGACTATGAAATGACTTTAAGTGGAGTTCGCGAGAGAGAAC 2031  
1973 |  
1981 GCTGACGCTGTGAAAGAGAGCTCTCATGTGGAGCTCAAAATGATGATGACCCACCTGGGA 2040  
1982 |  
2032 GCTGACGCTGTGAAAGAGAGCTCTCATGTGGAGCTCAAAATGATGATGACCCACCTGGGA 2091  
2033 |  
2041 CACCATGACACATCGTGAATCTGCTGGGGCATGACACTGTCAGGGCGGATGATGATG 2100  
2042 |  
2092 CACCATGACACATCGTGAATCTGCTGGGGCATGACACTGTCAGGGCGGATGATGATG 2151  
2093 |  
2101 ATTTTGAATATTTGTTGCTATGTTGCTCTCAACTACTAAAGATGAAAGAGAGAGAG 2160  
2102 |  
2152 ATTTTGAATATTTGTTGCTATGTTGCTCTCAACTACTAAAGATGAAAGAGAGAGAG 2211  
2153 |  
2161 TTTTCAAGGACATGAGAGAGATTTTAAAGGAAACATAATTTTCACTTCACTTCACTTTC 2220  
2162 |  
2212 TTTTCAAGGACATGAGAGAGATTTTAAAGGAAACATAATTTTCACTTCACTTCACTTTC 2271  
2213 |  
2221 CAGGACATTTCAATTTCCAGCATGCTGGTTCACGAGAGTTTCACTTCACTTCACTTTC 2280  
2222 |  
2272 CAGGACATTTCAATTTCCAGCATGCTGGTTCACGAGAGTTTCACTTCACTTCACTTTC 2331  
2273 |  
2281 GATCAGCTCTCAGGTTTCAATTTGGGAAATTTCAATTTCTGAGATGAGATTTCAATTTGA 2340  
2282 |  
2332 GATCAGCTCTCAGGTTTCAATTTGGGAAATTTCAATTTCTGAGATGAGATTTCAATTTGA 2391  
2333 |  
2341 AACCAGAGAGGCTGGCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2400  
2342 |

Db 2392 AACACAGAGAGGCTGGCAGAGAAGAGGAGGAAGATTGAACTGCTGACGTTTGAAGAC 2451  
Qy 2401 CTCCTTTGCTTTGCTTACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAAGTCGTGT 2460  
Db 2452 CTCCTTTGCTTTGCTTACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAAGTCGTGT 2511  
Qy 2461 GTCCACAGAGACTGGCAGCAGCAATGTGCTCACCCACGGGAAGTGGTGAAGATC 2520  
Db 2512 GTCCACAGAGACTGGCAGCAGCAATGTGCTCACCCACGGGAAGTGGTGAAGATC 2571  
Qy 2521 TGTGACTTTGGACTGCCCCGAGACATCCTGAGCGACTCCAGACTACGTGCTCAGGGGCAAC 2580  
Db 2572 TGTGACTTTGGACTGCCCCGAGACATCCTGAGCGACTCCAGACTACGTGCTCAGGGGCAAC 2631  
Qy 2581 GCACGCTCGCGTGAAGTGGATGGCACCAGAGACTTATTTGAAGGATCTACAATC 2640  
Db 2632 GCACGCTCGCGTGAAGTGGATGGCACCAGAGACTTATTTGAAGGATCTACAATC 2691  
Qy 2641 AAGAGTGAGCTGTGCTTACGGCATCTTCTCTGGGAGATATTTTCACTGGGTGTAAC 2700  
Db 2692 AAGAGTGAGCTGTGCTTACGGCATCTTCTCTGGGAGATATTTTCACTGGGTGTAAC 2751  
Qy 2701 CTTTACCTTGGCATCTTCTGTCAGCGCTAACTTCTATAAACTGATTCAGAGTGAATTTAA 2760  
Db 2752 CTTTACCTTGGCATCTTCTGTCAGCGCTAACTTCTATAAACTGATTCAGAGTGAATTTAA 2811  
Qy 2761 ATGGAGCAGCCATTTATGCCACAGAGGATATATTTTGAATGAATCACTGCTGGGT 2820  
Db 2812 ATGGAGCAGCCATTTATGCCACAGAGGATATATTTTGAATGAATCACTGCTGGGT 2871  
Qy 2821 TTTGACTCAAGAGCGGCCATCTTCCCCACCTGACTTCAATTTTAGGATGTACGCTG 2880  
Db 2872 TTTGACTCAAGAGCGGCCATCTTCCCCACCTGACTTCAATTTTAGGATGTACGCTG 2931  
Qy 2881 GCAGAGGCGAAGAAAGC-----ATGATCAGAAATCCATCC 2917  
Db 2932 GCAGAGGCGAAGAGCGATGTATCAGACATGGGTGGCAACGTCCAGAAATCCATCC 2991  
Qy 2918 ATCTACCAAAACAGCGGCCCTTACAGAGAGGGGCTCAGAGCC-CAGTGGCCACAG 2976  
Db 2992 ATCTACCAAAACAGCGGCCCTTACAGAGAGGGGCTCAGAGCGCCATCGCCACAG 3051  
Qy 2977 CGCAGGTGAAGATTCACAGAGAAAGTTCAGGAGGGGCTTGGACCCGCCACCT 3036  
Db 3052 GCCCAGGTGAAGATTCACAGAGAAAGTTCAGGAGGGGCTTGGACCCGCCACCT 3111  
Qy 3037 AGCAGGCTGTAGACCGCAGAGCAAGATTAGCTCGCTCTCAGGAAGCGCCCTACAGCG 3096  
Db 3112 AGCAGGCTGTAGACCGCAGAGCAAGATTAGCTCGCTCTCAGGAAGCGCCCTACAGCG 3171  
Qy 3097 CGTGTCTCGCTGGACTTTTCTAGATGCTGTCTGCCAATTAATCCAAAGTCACTTCTAT 3156  
Db 3172 CGTGTCTCGCTGGACTTTTCTAGATGCTGTCTGCCAATTAATCCAAAGTCACTTCTAT 3231  
Qy 3157 AAATCAACCTCTCTCGCAGCGGGGAGAGCCAAATATGAGACTTGTGGTAGGCC 3216  
Db 3232 AAATCAACCTCTCTCGCAGCGGGGAGAGCCAAATATGAGACTTGTGGTAGGCC 3291  
Qy 3217 GCCTACCTCTGGGGCTTT-----CCACGAGCTTGGGGGAAGCCATGATCTGAA 3268  
Db 3292 GCCTACCTCTGGGGCTTTCCAGGCCCTCCAGGCTTGGGGGAAGCCATGATCTGAA 3351  
Qy 3269 ATATAGTATATTTTGTGTAATACGTGAAACAAACCAAAACCCGTTTTTGTCTAAGGGAAG 3328  
Db 3352 ATATAGTATATTTTGTGTAATACGTGAAACAAACCAAAACCCGTTTTTGTCTAAGGGAAG 3411  
Qy 3329 CTAATATGATTTTAAATCTATGTTTAAATACATGTAATCTTTTCTATTTA 3389  
Db 3412 CTAATATGATTTTAAATCTATGTTTAAATACATGTAATCTTTTCTATTTA 3471  
Qy 3389 GTGATATATTTTATGGATGGAATAAACTTTCTACTGTAAATAAAAAA 3438  
Db 3472 GTGATATATTTTATGGATGGAATAAACTTTCTACTGTAGAAAAA 3521

## RESULT 12

US-08-434-878-1  
; Sequence 1, Application US/08434878  
; Patent No. 5957865  
; GENERAL INFORMATION:  
; APPLICANT: Bennett, Brian D.  
; APPLICANT: Broz, Susan D.  
; APPLICANT: Matthews, William  
; APPLICANT: Zeigler, Francis C.  
; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES THE  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/434,878  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hasak, Janet E.  
; REGISTRATION NUMBER: 28,616  
; REFERENCE/DOCKET NUMBER: 879  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-1896  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3521 bases  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-434-878-1

Query Match 96.8%; Score 3344; DB 2; Length 3521;  
Best Local Similarity 98.5%; Pred. No. 0;  
Matches 3418; Conservative 0; Mismatches 20; Indels 32; Gaps 3;

Qy 1 CGGCGCTGGCTACCGCGCTCCGAGGCCATCGGGGCTTGGCGCAGCGCAGCGACCGG 60  
Db 52 CGGCGCTGGCTACCGCGCTCCGAGGCCATCGGGGCTTGGCGCAGCGCAGCGACCGG 111  
Qy 61 CGGCTGCTGCTGCTTGTGTTTGTCTAGTATGTTCTTGTAGCCGTTTACAAACCAAGAC 120  
Db 112 CGGCTGCTGCTGCTTGTGTTTGTCTAGTATGTTCTTGTAGACCGTTTACAAACCAAGAC 171  
Qy 121 CTGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAAATGGCTCATCAGCGGGAAG 180  
Db 172 CTGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAAATGGCTCATCAGCGGGAAG 231  
Qy 181 CCATCATCTGATCCGAATGCTCGAGGATCCCGAAGACCTCCAGTGTACCCCGAGGCGC 240  
Db 232 CCATCATCTGATCCGAATGCTCGAGGATCCCGAAGACCTCCAGTGTACCCCGAGGCGC 291  
Qy 241 CAGAGTGAAGGACGCTATATGAAGCGCCACCGTGGAGGTGGCGGATCTGGGTCCATC 300  
Db 292 CAGAGTGAAGGACGCTATATGAAGCGCCACCGTGGAGGTGGCGGATCTGGGTCCATC 351  
Qy 301 ACCCTGCAAGTGCAGCTCGCCACCCAGGGGACCTTTCTGCTCTGGGTCTTTTAAGCAC 360



|||||  
352 ACCCTGCAAGTGCAGCTCGCCACCCAGGGACCTTTCTGCTCTGGGTCTTTAAGCAC 411  
Qy  
361 AGCTCCCTGGGCTGCCAGCCGCACTTTGATTTACAAACAGAGGAATCGTTTCCATGCCC 420  
Db  
412 AGCTCCCTGGGCTGCCAGCCGCACTTTGATTTACAAACAGAGGAATCGTTTCCATGCCC 471  
Qy  
421 ATCTTGAAACGTGACAGAGACCCAGCAGGAGAAATACCTTACCTTATTTACAGAGCGAAGCC 480  
Db  
472 ATCTTGAAACGTGACAGAGACCCAGCAGGAGAAATACCTTATTTACAGAGCGAAGCC 531  
Qy  
481 GCCAACTACACAGTACTGTTTACAGTGAATGTAAAGAGATACAGAGCTGTATGTGCTAAGG 540  
Db  
532 GCCAACTACACAGTACTGTTTACAGTGAATGTAAAGAGATACAGAGCTGTATGTGCTAAGG 591  
Qy  
541 AGACCTTACTTTAGAGATGGAACCCAGAGTGCATGCTCTGCATCTCCGAGGGTGT 600  
Db  
592 AGACCTTACTTTAGAGATGGAACCCAGAGTGCATGCTCTGCATCTCCGAGGGTGT 651  
Qy  
601 CCGGAGCCCACTGTGGAGTGGGTGCTCTGCAGCTCCACAGGGAAGCTGTAAAGAGAA 660  
Db  
652 CCGGAGCCCACTGTGGAGTGGGTGCTCTGCAGCTCCACAGGGAAGCTGTAAAGAGAA 711  
Qy  
661 GGCCCTGCTGTGTGAGAAAGGAGAAAGGTACTTCAATGATTTGTTCCGAAACAGACATC 720  
Db  
712 GGCCCTGCTGTGTGAGAAAGGAGAAAGGTACTTCAATGATTTGTTCCGAAACAGACATC 771  
Qy  
721 AGATGCTGTGTAGAAATGCACTGGGCGCGAATGCACCAAGCTGTTCCACATAGATCTA 780  
Db  
772 AGATGCTGTGTAGAAATGCACTGGGCGCGAATGCACCAAGCTGTTCCACATAGATCTA 831  
Qy  
781 AACACAGGCTCTCAGAGACACTGCCAGTATTTCTGAAAGTGGGGGAAACCTTTGTGG 840  
Db  
832 AACACAGGCTCTCAGAGACACTGCCAGTATTTCTGAAAGTGGGGGAAACCTTTGTGG 891  
Qy  
841 ATCAGGTGTAAAGCCATCCATGTGAACCAATGATTTGCGGCTCACCCTGGAGCTGGAAGAC 900  
Db  
892 ATCAGGTGTAAAGCCATCCATGTGAACCAATGATTTGCGGCTCACCCTGGAGCTGGAAGAC 951  
Qy  
901 AAAGCCCTGAGAGGGGAGCTACTTTGAGATGAGTACCTACTCCACAAACAGGACCAATG 960  
Db  
952 AAAGCCCTGAGAGGGGAGCTACTTTGAGATGAGTACCTACTCCACAAACAGGACCAATG 1011  
Qy  
961 ATTCCGATCTCTTGGCTTGTGCTTCCCTGGGAGGAGACGACCCGGATATTACACC 1020  
Db  
1012 ATTCCGATCTCTTGGCTTGTGCTTCCCTGGGAGGAGACGACCCGGATATTACACC 1071  
Qy  
1021 TGCTCTTCTCAAAGCACCCAGCCAGTCAGCGTTGGTGACCATCTAGAAAAAGGGTTT 1080  
Db  
1072 TGCTCTTCTCAAAGCACCCAGCCAGTCAGCGTTGGTGACCATCTAGAAAAAGGGTTT 1131  
Qy  
1081 ATAAACGCTACAGCTCGCAAGAGAGATGAAATTTGACCCGTAAGAAAGTTCTGCTTC 1140  
Db  
1132 ATAAACGCTACAGCTCGCAAGAGAGATGAAATTTGACCCGTAAGAAAGTTCTGCTTC 1191  
Qy  
1141 TCAGTCAGGTTTAAAGCGTACCCAGATCCGATCCGATCGATGATCTCTCAAGCTCA 1200  
Db  
1192 TCAGTCAGGTTTAAAGCGTACCCAGATCCGATCCGATCGATGATCTCTCAAGCTCA 1251  
Qy  
1201 TTTCTTGTGAAACAGAGAGGCTCGAGGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260  
Db  
1252 TTTCTTGTGAAACAGAGAGGCTCGAGGATGGGTACAGCATATCTAAATTTTGGCATCAT 1311  
Qy  
1261 AAGAAACAGCCAGGAGATACATATTTCTATCAGAAAAATGATGCGCCAGTTCCACCAA 1320  
Db  
1312 AAGAAACAGCCAGGAGATACATATTTCTATCAGAAAAATGATGCGCCAGTTCCACCAA 1371  
Qy  
1321 ATGTTACGCTGAATATTAAGAAAGAAACCTCAAGTGTAGCAATATGCTCAGCGCCAG 1380  
Db  
1372 ATGTTACGCTGAATATTAAGAAAGAAACCTCAAGTGTAGCAATATGCTCAGCGCCAG 1431  
Qy  
1381 GCGTCTGCTCTGATGGCTACCCGCTACCTCTGCACTGGAAGAGTGTTCGGAC 1440  
|||||

Db  
1432 GCGTCTGTTCTCTGATGGCTACCCGCTACCTCTTTGGACCTGGAAGAGTGTTCGGAC 1491  
Qy  
1441 AAATCTCCCAATTCGACGAGGAATCCAGAGAGTTCGAAATAAAAAAGCTTAACAGA 1500  
Db  
1492 AAATCTCCCAATTCGACGAGGAATCCAGAGAGTTCGAAATAAAAAAGCTTAACAGA 1551  
Qy  
1501 AAAGTGTTTGGCCAGTGGGTGTCGAGCAGTACTCTTAAATATGAGTGGAGCCGGGAAAGG 1560  
Db  
1552 AAAGTGTTTGGCCAGTGGGTGTCGAGCAGTACTCTTAAATATGAGTGGAGCCGGGAAAGG 1611  
Qy  
1561 CTTCTGTGTAATGCTGTGCGTACAAATTCATATGGGCAAGTCTTATGCGAAACCAATCTTTTA 1620  
Db  
1612 CTTCTGTGTAATGCTGTGCGTACAAATTCATATGGGCAAGTCTTATGCGAAACCAATCTTTTA 1671  
Qy  
1621 AACTCACAGCCCTTCCCTTTTCATCCAGACAACTCTCTTATGCGACCAATTTGG 1680  
Db  
1672 AACTCACAGCCCTTCCCTTTTCATCCAGACAACTCTCTTATGCGACCAATTTGG 1731  
Qy  
1681 CTTCTGTCTCCCTTTCAITTTGTTTCTCATTTGTTGATCTGCCACAAATACAAAAAGCAA 1740  
Db  
1732 CTTCTGTCTCCCTTTCAITTTGTTTCTCATTTGTTGATCTGCCACAAATACAAAAAGCAA 1791  
Qy  
1741 TTTAGGTACGAGAGTCAAGTGCAGATGATCCAGTGCAGTGGCCCTCGGATTAACAGATAC 1800  
Db  
1792 TTTAGGTACGAGAGTCAAGTGCAGATGATCCAGTGCAGTGGCCCTCGGATTAACAGATAC 1851  
Qy  
1801 TTTCTACGTTGACTTACGGGACTATGATGACCTTAAGTGGGAGTTCGCGAGAGAGAC 1860  
Db  
1852 TTTCTACGTTGACTTACGGGACTATGATGACCTTAAGTGGGAGTTCGCGAGAGAGAC 1911  
Qy  
1861 TTAGAGTTTGGGAAAGTCTGGGGTCTGGCGCTTTTCGGGAGGTCATGAACCCACGGCC 1920  
Db  
1912 TTAGAGTTTGGGAAAGTCTGGGGTCTGGCGCTTTTCGGGAGGTCATGAACCCACGGCC 1971  
Qy  
1921 TATGGCAATTAGTAAAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAAAGAGAA 1980  
Db  
1972 TATGGCAATTAGTAAAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAAAGAGAA 2031  
Qy  
1981 GCTGACAGCTGTGAAAAAGAGCTCTCATGTGCGAGCTCAAAAATGATGACCCACCTGGGA 2040  
Db  
2032 GCTGACAGCTGTGAAAAAGAGCTCTCATGTGCGAGCTCAAAAATGATGACCCACCTGGGA 2091  
Qy  
2041 CACCATGACACATCGTGAATCTGCTGGGGCATGCACATGTGAGGGCCAGTGTACTTTG 2100  
Db  
2092 CACCATGACACATCGTGAATCTGCTGGGGCATGCACATGTGAGGGCCAGTGTACTTTG 2151  
Qy  
2101 ATTTTGAATATTGTTGCTATGTTGACCTCTCAACTACCTAAAGAAAGTAAAGAGAGAGAG 2160  
Db  
2152 ATTTTGAATATTGTTGCTATGTTGACCTCTCAACTACCTAAAGAAAGTAAAGAGAGAGAG 2211  
Qy  
2161 TTTTCAAGGACATGGACAGAGATTTTAAAGAAATATTTTCAATTTTCAATTTTCACTTTTC 2220  
Db  
2212 TTTTCAAGGACATGGACAGAGATTTTAAAGAAATATTTTCAATTTTCACTTTTCACTTTTC 2271  
Qy  
2221 CAGGACATTTCAAAATTCAGCATGCTGGTTTCAGGAGAGTTCAGTTACACCCGCTTTG 2280  
Db  
2272 CAGGACATTTCAAAATTCAGCATGCTGGTTTCAGGAGAGTTCAGTTACACCCGCTTTG 2331  
Qy  
2281 GATCAGCTCTCAGGGTTCAATGGGAATTTCAATTTCAATTTCTGAAGATGAGATTAATGAA 2340  
Db  
2332 GATCAGCTCTCAGGGTTCAATGGGAATTTCAATTTCAATTTCTGAAGATGAGATTAATGAA 2391  
Qy  
2341 AACAGAGAGGCTGGCAGAGAGAGAGAGAGAGATTTGAAACGTCGTGACGTTTGAAGAC 2400  
Db  
2392 AACAGAGAGGCTGGCAGAGAGAGAGAGAGAGATTTGAAACGTCGTGACGTTTGAAGAC 2451  
Qy  
2401 CTCCTTTGCTTTGCTTACCAAGTGGCCAGAGGATGAAATTCCTTGGAGTTCAAATCTGTT 2460  
Db  
2452 CTCCTTTGCTTTGCTTACCAAGTGGCCAGAGGATGAAATTCCTTGGAGTTCAAATCTGTT 2511  
Qy  
2461 GTCCACAGAGACCTGGCAG 2520  
Db  
2512 GTCCACAGAGACCTGGCAG 2571

```
QY 2521 TGTGACTTTGGACTGGCCGAGACATCTTGAGGACTCCAGCTACGTCTCAGGGGCAAC 2580
Db |||||||
QY 2572 TGTGACTTTGGACTGGCCGAGACATCTTGAGGACTCCAGCTACGTCTCAGGGGCAAC 2631
Db |||||||
QY 2581 GCA CGGCTGCCGCTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGGATCTACACAATC 2640
Db GCACGGCTGCCGCTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGGATCTACACAATC 2691
QY 2641 AAGAGTGAGCTGTGCTCTACCGGATCTCTCTCTGGAGATATTTTCACTGGGTGTGAAC 2700
Db |||||||
QY 2692 AAGAGTGAGCTGTGCTCTACCGGATCTCTCTCTGGAGATATTTTCACTGGGTGTGAAC 2751
QY 2701 CTTTACCTTGGCATCTCTGTCGACGCTACTTCTATAAATCTAATCAGAGTGGATTTAA 2760
Db |||||||
QY 2752 CTTTACCTTGGCATCTCTGTCGACGCTACTTCTATAAATCTAATCAGAGTGGATTTAA 2811
QY 2761 ATGGAGCAGCCATTTCTATGCCACAGAGGGATATCTTTGTAATGCAATCTCTGCTGGGCT 2820
Db |||||||
QY 2812 ATGGAGCAGCCATTTCTATGCCACAGAGGGATATCTTTGTAATGCAATCTCTGCTGGGCT 2871
QY 2821 TTTGACTCAAGAGAGCGGCTATCTCTTCCCACTGACTTCAATTTTATAGGATGTGAGCTG 2880
Db |||||||
QY 2872 TTTGACTCAAGAGAGCGGCTATCTCTTCCCACTGACTTCAATTTTATAGGATGTGAGCTG 2931
QY 2881 GCAGAGGAGAGAGAGC-----ATGATATCAGAAATCCATCC 2917
Db |||||||
QY 2932 GCAGAGGAGAGAGAGAGAGATGATATCAGAAATCTGAGGAGGCTTTGGACCCCGCACCT 2991
QY 2918 ATCTACCAAAACAGCGGCCCTCTAGCAGAGAGGCGGCTCAGAGCC-CAGTCGCCACAG 2976
Db |||||||
QY 2992 ATCTACCAAAACAGCGGCCCTCTAGCAGAGAGGCGGCTCAGAGCCCGCATCGCCACAG 3051
QY 2977 CGCAGAGTGAAATCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3036
Db |||||||
QY 3052 GCCCAGGTGAAATCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3111
QY 3037 AGCAGGCTGTAGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3096
Db |||||||
QY 3112 AGCAGGCTGTAGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3171
QY 3097 CGTTGCTTGGCTGACTTTTCTCTAGATGCTGCTGCCATTAATCTCCAAAGTGACTTCTAT 3156
Db |||||||
QY 3172 CGTTGCTTGGCTGACTTTTCTCTAGATGCTGCTGCCATTAATCTCCAAAGTGACTTCTAT 3231
QY 3157 AAAATCAAACTCTCTCGCACAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3216
Db |||||||
QY 3232 AAAATCAAACTCTCTCGCACAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3291
QY 3217 GCCTACCTGGGGGCTTT-----CCACGAGCTTGAGGGGAAAGCCATGTATCTGAA 3268
Db GCCTACCTGGGGGCTTTTTCAGGCCCCCAGGCTTGAGGGGAAAGCCATGTATCTGAA 3351
QY 3269 ATATAGTATATTTTGTAAATACGTGAAACAAACCAACCGGTTTGTCTAAGGGAAG 3328
Db |||||||
QY 3352 ATATAGTATATTTTGTAAATACGTGAAACAAACCAACCGGTTTGTCTAAGGGAAG 3411
QY 3329 CTAATATGATTTTAAATATCTATGTTTAAATATCTATGTAATCTTTTTCATCTATTTA 3388
Db |||||||
QY 3412 CTAATATGATTTTAAATATCTATGTTTAAATATCTATGTAATCTTTTTCATCTATTTA 3471
QY 3389 GTGATATATTTTATGATGGAATAAATCTTTCTACTGTAAATAAAAAA 3438
Db |||||||
QY 3472 GTGATATATTTTATGATGGAATAAATCTTTCTACTGTAAATAAAAAA 3521
```

## RESULT 13

PCT-US95-03718-1

; Sequence 1, Application PC/TUS9503718

; GENERAL INFORMATION:

; APPLICANT: GENENTECH, INC.

; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES THEREOF

; NUMBER OF SEQUENCES: 4

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/03718
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wendy M. Lee
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: 879PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3521 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; PCT-US95-03718-1
```

Query Match 96.8%; Score 3344; DB 5; Length 3521;

Best Local Similarity 98.5%; Pred. No. 0;

Matches 3418; Conservative 0; Mismatches 20; Indels 32; Gaps 3;

```
QY 1 GCGGCTGGCTTACCGCGCTCCGGAGGCCATCGGGCGTTGGCGCAGCGAGCGACCGG 60
Db 52 GCGGCTGGCTTACCGCGCTCCGGAGGCCATCGGGCGTTGGCGCAGCGACCGG 111
QY 61 CGGCTGCTGCTGCTTGTGTTTGTCAAGTATGATTTCTGAGACCGTTTACAAACCAAGAC 120
Db 112 CGGCTGCTGCTGCTTGTGTTTGTCAAGTATGATTTCTGAGACCGTTTACAAACCAAGAC 171
QY 121 CTGCTGTGATCAAGTGTGTTTAAATCAGTCATCAGAAACAAATGGCTCATCAGCGGGAAG 180
Db 172 CTGCTGTGATCAAGTGTGTTTAAATCAGTCATCAGAAACAAATGGCTCATCAGCGGGAAG 231
QY 181 CCATCATCTGATACCGAATGTCGAGGATCCCGAGAACCTCCAGTGTACCCGAGGCGC 240
Db 232 CCATCATCTGATACCGAATGTCGAGGATCCCGAGAACCTCCAGTGTACCCGAGGCGC 291
QY 241 CAGAGTGAAAGGAGCGGTATATGAAAGCGGCCA CCGTGAGGTGGCGGAGTCTGGGTCCATC 300
Db 292 CAGAGTGAAAGGAGCGGTATATGAAAGCGGCCA CCGTGAGGTGGCGGAGTCTGGGTCCATC 351
QY 301 ACCCTGCAAGTGCAAGTCTCGCCA CCCCAGGGGACCTTTCTGCTCTGGGTCTTTAAGCAC 360
Db 352 ACCCTGCAAGTGCAAGTCTCGCCA CCCCAGGGGACCTTTCTGCTCTGGGTCTTTAAGCAC 411
QY 361 AGCTCCCTGGCTGCCAGCGGCTTTGATTTTACAAACAGAGAAATCGTTTCCATGGCC 420
Db 412 AGCTCCCTGGCTGCCAGCGGCTTTGATTTTACAAACAGAGAAATCGTTTCCATGGCC 471
QY 421 ATCTTGAACGTGACAGAGACCCAGGAGGAGATACCTACTCCATATTTACAGAGCGACGC 480
Db 472 ATCTTGAACGTGACAGAGACCCAGGAGGAGATACCTACTCCATATTTACAGAGCGACGC 531
QY 481 GCCAACTACAGTACTGTTTACAGTGAATGTGAAGAGATACACAGCTGTATGTGCTAAGG 540
|||||
```

Db 532 GCCAACTACACAGTACTGTTCCACAGTGAATGTAAGAGATACACAGCTGTATGTGCTAAGG 591  
Qy 541 AGACCTTACTTTAGGAAGATGGAAAA CCAGATGCACTGCTCTGCAATCTCCGAGGGTGT 600  
Db 592 AGACCTTTACTTTAGGAAGATGGAAAA CCAGATGCACTGCTCTGCAATCTCCGAGGGTGT 651  
Qy 601 CCGGAGCCCACTGTGGAGTGGTGTCTGCAAGTCCCAACAGGAAGAGCTGTAAGAAGAA 660  
Db 652 CCGGAGCCCACTGTGGAGTGGTGTCTGCAAGTCCCAACAGGAAGAGCTGTAAGAAGAA 711  
Qy 661 GGCCCTGCTGTGTGTCAGAAAAGGAGAAAAGGTACTTTCAATGAGTGTTCGGAAACAGACATC 720  
Db 712 GGCCCTGCTGTGTGTCAGAAAAGGAGAAAAGGTACTTTCAATGAGTGTTCGGAAACAGACATC 771  
Qy 721 AGATGCTGTGTAGAAATGCACTGGGCGCGAATGCAACCAAGCTGTTCACCATGATCTA 780  
Db 772 AGATGCTGTGTAGAAATGCACTGGGCGCGAATGCAACCAAGCTGTTCACCATGATCTA 831  
Qy 781 AACAGGCTCTCAGAGCACACTGCCCGAGTTATTCCTGAAAGTGGGGGAAACCCCTTGTGG 840  
Db 832 AACAGGCTCTCAGAGCACACTGCCCGAGTTATTCCTGAAAGTGGGGGAAACCCCTTGTGG 891  
Qy 841 ATCAGGTGTAAAGGCATCTCATGTGAACCATGCAATTCGGGCTCACCTGGGAGCTGGAAGAC 900  
Db 892 ATCAGGTGTAAAGGCATCTCATGTGAACCATGCAATTCGGGCTCACCTGGGAGCTGGAAGAC 951  
Qy 901 AAAGCCCTGGAGGGGAGGAGCTACTTTGAGATGAGTACTCTCCACAAAACAGGACCATG 960  
Db 952 AAAGCCCTGGAGGGGAGGAGCTACTTTGAGATGAGTACTCTCCACAAAACAGGACCATG 1011  
Qy 961 ATTCCGATTTCTTTGGGCTTTGTGTCTTCCGTGGGAGGAAGACGACCGGATATTACACC 1020  
Db 1012 ATTCCGATTTCTTTGGGCTTTGTGTCTTCCGTGGGAGGAACGACACCGGATATTACACC 1071  
Qy 1021 TGCTTTCTCTCAAGAGCCCCAGCCAGTCAGCGTTGTGTGACCATCTCTAGAAAGAGGTTT 1080  
Db 1072 TGCTTTCTCTCAAGAGCCCCAGCCAGTCAGCGTTGTGTGACCATCTCTAGAAAGAGGTTT 1131  
Qy 1081 ATAAAGCTACAGCTCGCAAGAGAGATGATGAATTTGACCCGTACGAAAAGTTCTGCTTC 1140  
Db 1132 ATAAAGCTACAGCTCGCAAGAGAGATGATGAATTTGACCCGTACGAAAAGTTCTGCTTC 1191  
Qy 1141 TCAGTCAAGTTTAAAGGTACCAAGAACTCGATGCGAGTGGATCTTCTCTCAAGCCCTCA 1200  
Db 1192 TCAGTCAAGTTTAAAGGTACCAAGAACTCGATGCGAGTGGATCTTCTCTCAAGCCCTCA 1251  
Qy 1201 TTTTCTTTGTGAACAGAGAGGCTCGAGGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260  
Db 1252 TTTTCTTTGTGAACAGAGAGGCTCGAGGATGGGTACAGCATATCTAAATTTTGGCATCAT 1311  
Qy 1261 AAGAAACAGCCAGAGAGTACATATTTCTATGCAGAAAATGATGACGCCAGTTTCAACAAA 1320  
Db 1312 AAGAAACAGCCAGAGAGTACATATTTCTATGCAGAAAATGATGACGCCAGTTTCAACAAA 1371  
Qy 1321 ATGTTCAAGTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCTTCAGCCAGCCAG 1380  
Db 1372 ATGTTCAAGTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCTTCAGCCAGCCAG 1431  
Qy 1381 GCGTCTCTTCTCTGATGCTACCCGCTACCCCTCTTGGACCTGGAGAGAGTGTTCGGAC 1440  
Db 1432 GCGTCTCTTCTCTGATGCTACCCGCTACCCCTCTTGGACCTGGAGAGAGTGTTCGGAC 1491  
Qy 1441 AAATCTCCCAATTCACGGAGGAAATCCCAAGAGAGTTTGGAAATAAAAGGCTTAACAGA 1500  
Db 1492 AAATCTCCCAATTCACGGAGGAAATCCCAAGAGAGTTTGGAAATAAAAGGCTTAACAGA 1551  
Qy 1501 AAAGTGTTCGCGAGTGGTGTGCGAGAGTACTCTAAATATGAGTGAAGCCGGGAAAGGG 1560  
Db 1552 AAAGTGTTCGCGAGTGGTGTGCGAGAGTACTCTAAATATGAGTGAAGCCGGGAAAGGG 1611  
Qy 1561 CTTCTGTGTAATGCTGTGGGTACAAATTTATAGGCGAGCTCTTGGGAAACCATCTTTTAA 1620  
Db 1612 CTTCTGTGTAATGCTGTGGGTACAAATTTATAGGCGAGCTCTTGGGAAACCATCTTTTAA 1671

Qy 1621 AACTCACAGGCCCCCTTCCCTTTTCATCCAGAGCAAACTCTCTTCTATGCGACCAATTGGG 1680  
Db 1672 AACTCACAGGCCCCCTTCCCTTTTCATCCAGAGCAAACTCTCTTCTATGCGACCAATTGGG 1731  
Qy 1681 CTCTGTCTCCCTTTCAATTTGTTTCTCATTTGTTGATCTGCCCAAAATACAAAAGCAA 1740  
Db 1732 CTCTGTCTCCCTTTCAATTTGTTTCTCATTTGTTGATCTGCCCAAAATACAAAAGCAA 1791  
Qy 1741 TTTAGGTACAGAGTCAAGCTGACGATGATCCAGGTGACTGCGCCCCCTCGATAAACAGGTAC 1800  
Db 1792 TTTAGGTACAGAGTCAAGCTGACGATGATCCAGGTGACTGCGCCCCCTCGATAAACAGGTAC 1851  
Qy 1801 TTCTACGTTGACTTCAGGGAATAATGAACTTAAGTGGGAGTTCCCGAGAGAGAAC 1860  
Db 1852 TTCTACGTTGACTTCAGGGAATAATGAACTTAAGTGGGAGTTCCCGAGAGAGAAC 1911  
Qy 1861 TTAGAGTTTCGGAGGCTCTGGGCTCTGGCGCTTTTCGGGAGGTTGATGAAACGCCACGCC 1920  
Db 1912 TTAGAGTTTCGGAGGCTCTGGGCTCTGGCGCTTTTCGGGAGGTTGATGAAACGCCACGCC 1971  
Qy 1921 TATGGCATTAAGTAAACCGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAACAGAGAA 1980  
Db 1972 TATGGCATTAAGTAAACCGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAACAGAGAA 2031  
Qy 1981 GCTGACAGCTGTGAAAAAGAAAGCTCTCATGTCCGAGCTCAAAATGATGACCCACCTGGGA 2040  
Db 2032 GCTGACAGCTGTGAAAAAGAAAGCTCTCATGTCCGAGCTCAAAATGATGACCCACCTGGGA 2091  
Qy 2041 CACCATGACAACTGTAATCTCTCTGGGGCATGCACTGTGAGGGCCAGTGTACTTG 2100  
Db 2092 CACCATGACAACTGTAATCTCTCTGGGGCATGCACTGTGAGGGCCAGTGTACTTG 2151  
Qy 2101 ATTTTTCGAATATGTTCTGATGTTGACCTCTCAACTACCTAAAGAGTAAAGAGAGAG 2160  
Db 2152 ATTTTTCGAATATGTTCTGATGTTGACCTCTCAACTACCTAAAGAGTAAAGAGAGAG 2211  
Qy 2161 TTTTCAAGGACATGGAACAGAGATTTTAAAGGAAATATTTTCACTTCTTACCTACTTTTC 2220  
Db 2212 TTTTCAAGGACATGGAACAGAGATTTTAAAGGAAATATTTTCACTTCTTACCTACTTTTC 2271  
Qy 2221 CAGGCAATTCAAATTCAGCATGCTGTTTACAGAGATTTTCACTTCTTACCGGCCCTTG 2280  
Db 2272 CAGGCAATTCAAATTCAGCATGCTGTTTACAGAGATTTTCACTTCTTACCGGCCCTTG 2331  
Qy 2281 GATCAGCTCTCAGGTTCAATGGGAATCAATTCATCTGAAGATGAGATTGAATATCAA 2340  
Db 2332 GATCAGCTCTCAGGTTCAATGGGAATTTCAATTTCAITCTGAAGATGAGATTGAATATCAA 2391  
Qy 2341 AACCAAGAGGCTTGGCAGAGAGAGAGAGATTTGAACGTGCTGACGTTTGAAGAC 2400  
Db 2392 AACCAAGAGGCTTGGCAGAGAGAGAGAGATTTGAACGTGCTGACGTTTGAAGAC 2451  
Qy 2401 CTCCTTTGCTTTGCTTACCAAGTGGCCAAAGGATGAAATTTCTTGGAGTTCAAGTCGTGT 2460  
Db 2452 CTCCTTTGCTTTGCTTACCAAGTGGCCAAAGGATGAAATTTCTTGGAGTTCAAGTCGTGT 2511  
Qy 2461 GTCCACAGAGACCTTGGCAGGAGGAGATGTTGTTGTTCAACCAAGGAGGAGTGGTGAAGATC 2520  
Db 2512 GTCCACAGAGACCTTGGCAGGAGGAGATGTTGTTGTTCAACCAAGGAGGAGTGGTGAAGATC 2571  
Qy 2521 TGTGACTTTTGGACTGGCCCGAGACATCTCGAGGAGCTCCAGTACGTCGTGAGGGGCAAC 2580  
Db 2572 TGTGACTTTTGGACTGGCCCGAGACATCTCGAGGAGCTCCAGTACGTCGTGAGGGGCAAC 2631  
Qy 2581 GCACGGCTGCGGTGAGTGGATGGCACCCGAGAGCTTATTTGAAGGGATCTACACAAATC 2640  
Db 2632 GCACGGCTGCGGTGAGTGGATGGCACCCGAGAGCTTATTTGAAGGGATCTACACAAATC 2691  
Qy 2641 AAGAGTGAAGCTGTGGTCTTAAGGATCCCTTCTCTGGAGATATTTTCACTGGGTTGTGAAC 2700  
Db 2692 AAGAGTGAAGCTGTGGTCTTAAGGATCCCTTCTCTGGAGATATTTTCACTGGGTTGTGAAC 2751

2701 CCTTACCTGGCATTCTGTGACGCTAACTTCTATATAAAGTGAATCAGAGTGGAATTA 2760  
2752 CCTTACCTGGCATTCTGTGACGCTAACTTCTATATAAAGTGAATCAGAGTGGAATTA 2811  
2761 ATGGAGCAGCCATTCTATGCCAGAGAGGATATCTTTGTAATCAATCCCTGCTGGCT 2820  
2812 ATGGAGCAGCCATTCTATGCCAGAGAGGATATCTTTGTAATCAATCCCTGCTGGCT 2871  
2821 TTTGACTCAAGAGAGGCGCATCTTCCCAACCTGACTTCTATTTAGGATGTGAGCTG 2880  
2872 TTTGACTCAAGAGAGGCGCATCTTCCCAACCTGACTTCTATTTAGGATGTGAGCTG 2931  
2881 GCAGAGCAGAGAGAGC-----ATGTATCAGAAATCATCATCC 2917  
2932 GCAGAGCAGAGAGAGCAGATGATCAGAACATGGGTGGCAACGTCCCAAGATCATCATCC 2991  
2918 ATCTACCAAAAACAGCGCGCCCTCAGCAGAGAGGCGGCTCAGAGCC-CAGTCGCCACAG 2976  
2992 ATCTACCAAAAACAGCGCGCCCTCAGCAGAGAGGCGGCTCAGAGCGCGCATCGCCACAG 3051  
2977 CGCAGGTGAAGATTCAGAGAGAAAGATTAGCGAGAGGCGCTTGGACCCCGCCACCT 3036  
3052 GCCCAGGTGAAGATTCAGAGAGAAAGATTAGCGAGAGGCGCTTGGACCCCGCCACCT 3111  
3037 AGCAGGCTGTAGACCGCAGAGCAGAGATTAGCGAGAGGCGCTTGGACCCCGCCACCT 3096  
3112 AGCAGGCTGTAGACCGCAGAGCAGAGATTAGCGAGAGGCGCTTGGACCCCGCCACCT 3171  
3097 CGTTGCTTGGCTGAGCTTTTCTAGATGCTGTCTGCAATCTTCAAAAGTACTTCTAT 3156  
3172 CGTTGCTTGGCTGAGCTTTTCTAGATGCTGTCTGCAATCTTCAAAAGTACTTCTAT 3231  
3157 AAAATCAAACTCTCTCGCAGAGCGGAGGCAATTAATGAGACTTGTGTGAGGCC 3216  
3232 AAAATCAAACTCTCTCGCAGAGCGGAGGCAATTAATGAGACTTGTGTGAGGCC 3291  
3217 GCCTACCTCGGGGCTTT-----CCACGAGCTTGAGGGGAAAGCCATGATCTGAA 3268  
3292 GCCTACCTCGGGGCTTTCCAGGCCCGCCAGGCTTGGGGGAAAGCCATGATCTGAA 3351  
3269 ATATAGTATATTTCTGTAATACTGTAACCAACCAACCGTTTGTGTAAGGAAAG 3328  
3352 ATATAGTATATTTCTGTAATACTGTAACCAACCAACCGTTTGTGTAAGGAAAG 3411  
3329 CTAATATAGTATTTTAAATCTATGTTTAAATATCTATGTAATCTTTTCACTATT 3388  
3412 CTAATATAGTATTTTAAATCTATGTTTAAATATCTATGTAATCTTTTCACTATT 3471  
3389 GTGATATATTTTATGATGGAATAAATCTTCTACTGTAAAAA 3438  
3472 GTGATATATTTTATGATGGAATAAATCTTCTACTGTAAAAA 3521

RESULT 14  
US-07-977-451-3  
Sequence 3, Application US/07977451  
Patent No. 5270458  
GENERAL INFORMATION:  
APPLICANT: Lemischka, Ihor R.  
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESS: 180 Varick Street  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10014  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/977,451  
FILING DATE: 19921119  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US UNASSIGNED  
FILING DATE: 12-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/906,397  
FILING DATE: 26-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US PCT/US92/05401  
FILING DATE: 26-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: TW 81102961  
FILING DATE: 15-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US PCT/US92/02750  
FILING DATE: 02-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/813,593  
FILING DATE: 24-DEC-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/793,065  
FILING DATE: 15-NOV-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/728,913  
FILING DATE: 28-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Felt, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3501 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 58..3039  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 139..3036  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 58..138  
US-07-977-451-3

Query Match: 64.5%; Score 2227.2; DB 1; Length 3501;  
Best Local Similarity: 79.6%; Pred. No. 0;  
Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;  
QY 8 GGCTACCGCGCTCCGAGGCGCATCGGGCGTGGCGCAGCGCAGCGCGCGGCTGC 67  
DB 35 GGAGACCCCGGCTCCGAGGCGCATCGCGCGTGGC---GGCGACGGGGACCGTGC 91  
QY 68 TGCTGCTTGTGTTTGTGTCAGTAATGATTCCTTGGAGACCGTTACAAACCAAGACCTGCTG 127  
DB 92 CGCTGCTGCTGTTTGTGTTTCTGCAATGATATTTGGGACTATTACAAATCAAGATCTGCTG 151  
QY 128 TGATCAAGTGTGTTTATTAATCAATGAGAACATGAGGCTCATCAGCGGGGAAAGCCATCAT 187

Db 152 TGATCAAGTGTGTTTAAATCAATCAATAAGAAATGATTCATCAGTGGGAAAGTCATCAT 211  
Qy 188 CGTACCGATGTCGAGGATCCCAAGAGACCTCCAGTGTACCCCGAGGCGCCAGAGTG 247  
Db 212 CATATCCCATGTATCAGAAATCCCGGAAGACCTCGGGTGTGCGTTTGAGACCCAGAGCT 271  
Qy 248 AAGGACCGGTATATGAAGCGGCACCGTGGAGGTGGCCGAGTCTGGGTCCATCACCCTGC 307  
Db 272 CAGGACAGGTACAGAGCTGCGGTGTGGAAGTGAATGATCTGCTTCATCACTACACTGC 331  
Qy 308 AAGTCAGCTGCCACCCAGGGGACCTTTCTGCTCTGGGTCTTTAAGCAGCAGCTCCC 367  
Db 332 AAGTCTGCTCGATGCGCCAGGGAACATTTCTGCTCTGGGTCTTTAAGCAGCAGCTCCC 391  
Qy 368 TGGGCTGCCAGCGGCACATTTGATTTTACAAAAAGAGGAATGTTTCCATGGCCATCTGA 427  
Db 392 TGAATTGGCCAGCCACATTTTGAATTTTACAAAAAGAGGAGTGTGTTCCATGGTCATTTGA 451  
Qy 428 AGTGCACAGAGACCCAGGCGAGGAATACCTACTCCATATTCAGAGCGAAACGGCCAACT 487  
Db 452 AAATGACAGAAACCCAGCTGGAGAAATACCTACTTTTATTCAGAGTGAAGCTACCAAT 511  
Qy 488 ACACAGTACTGTCTCAAGTGAATGTAAGAGATACACAGCTGTATGTCTTAAGGAGACCTT 547  
Db 512 ACACAATATTTTACAGTGAATGAAGAAATACCTGCTTTACACATTTAAGAGACCTT 571  
Qy 548 ACTTTAGGAAGATGAAAAACAGGATGCACTGCTCTGCAATCTCCAGAGGTGTTCCGAGC 607  
Db 572 ACTTTAGAAAAATGAAAAACAGGACGCGCTGCTGCAATCTGAGAGCGTTCAGAGC 631  
Qy 608 CCATGTCGAGTGGGTGCTCTCAGCTCCCAAGGGAAGCTGTAAGAGAGAGCCCTG 667  
Db 632 CGATCGTGAATGGGTGCTTTGCGAATTCACAGGGGAAAGCTGTAAAGAGAAAGTCCAG 691  
Qy 668 CTGTGTCAGAAAGAGGAAAGTACTTCATGAGTTGTTTCGAAACAGACATCAGATGCT 727  
Db 692 CTGTGTTTAAAAAGAGGAAAGTGTCTCAATGAATTTATTTGGACGACATAAGTGTCT 751  
Qy 728 GTGCTAGAAATGCATCGGCGCGGAATGCAACAGCTGTCTCAATAGATCTPAAACAGG 787  
Db 752 GTGCCAGAAATGAATCTGGGCAAGGAATGCACCAGCTGTTCACATAGATCTPAAATCAA 811  
Qy 788 CTCCTCAGAGCACACTGCCAGTATTTCTGAAAGTGGGGAACCTTGTGGATCAGGT 847  
Db 812 CTCCTCAGACCACATTCGCCAATTTATTTCTTAAAGTAGGGAAACCTTTATGGATAAGGT 871  
Qy 848 GTAAGGCCATCCATGTGAACATGATTCGGGCTCACCTGGGAGCTGGAAGACAAGGCC 907  
Db 872 GCAAGCTGTTTCATGTGAACCATGATTCGGGCTCACCTGGGAATTTAGAAAAAANGCAC 931  
Qy 908 TGGAGGAGGCGAGCTPCTTTGAGATGAGTACCTACTCCACAAACAGGACCATGATTCGGA 967  
Db 932 TCGAGGAGGGCAA CTACTTTGAGATGAGTACCTATTCAACAAACAGAACTATGATACGGA 991  
Qy 968 TTCTCTGGCCCTTTGTCTCCGTTGGGAAGAAACGACACCGGATPATTTACACTGCTCT 1027  
Db 992 TTCTGTTTGTCTTTGTATCATCAGTGGCAAGAAACGACACCGGATPACTACACTTGTCT 1051  
Qy 1028 CCTCAAAGCACCCAGCGAGTCAGGTGTTGAGCACTCTAGAAAAGGGTTTATAAGC 1087  
Db 1052 CTTCAAAGCATCCCATGCTTGGTTGGTTTACCATCGTAGAAAGGGAATTTAATG 1111  
Qy 1088 CTACCAGCTCGAAGAGAGATGAAATTTGACCCGTAAGAAAAGTTCTGCTTCTCAGTCA 1147  
Db 1112 CTACCAATTTCAAGTGAAGATGATGAATTTGACCAATATGAAGAGTTTGTCTTCTGCA 1171  
Qy 1148 GGTTTAAGCGTACCCAGAAATCCGATGCA GTGGATCTTCTCTCAAGCTCATTTCTCT 1207  
Db 1172 GGTTTAAGCGTACCCAGAAATCAGATGTA GTGGACCTTCTCTGAAATCATTTCTCT 1231  
Qy 1208 GTGAACAGAGGCGCTGGAGATGGGTACAGCATATCTAAATTTTGGCATCATAGAACA 1267

Db 1232 GTGAGCAAAAGGGTCTTTGATAACGGATACAGCATATCAAGTTTGTCAATCATAAAGCAC 1291  
Qy 1268 AGCCAGGAGTACATATTTCTATGAGAAAATGATGACGCCCTAGTTTCACAAAATGTTCA 1327  
Db 1292 AGCCAGGAGAAATATATATTTCCATGCAAAAATGATGATGCCAAATTTACAAAATGTTCA 1351  
Qy 1328 CGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAATGCTCAGCCAGCCAGCGCTCT 1387  
Db 1352 CGCTGAATATAAGAGGAAACCTCAAGTGTGCGAAGCATCGGCAAGTCAGCGCTCT 1411  
Qy 1388 GTTCTCTGATGGCTACCCGCTACCCCTTTGGAACCTGGAAGAAAGTGTTCGACAAAATCTC 1447  
Db 1412 GTTCTCTGGATGATACCCATTTACCATCTTGACCTTGAAGAAAGTGTTCAGACAAGTCTC 1471  
Qy 1448 CCAATTCACGAGGAAATCCAGAAAGAGTGTGGAATAAAAGGCTTAAACAGAAAAGTGT 1507  
Db 1472 CCAACTGCACAGAGAGATCACAGAAAGAGTCTGGAATAGAAGGCTTAAACAGAAAAGTGT 1531  
Qy 1508 TTGGCCAGTGGGTGTCAGCAGTACTCTAAATATAGTAGTGGGCGGGAAGGGCTTCTGG 1567  
Db 1532 TTGGACAGTGGGTGTCAGCAGTACTCTAAACATGAGTGAAGCCATAAAAGGGTTCCTGG 1591  
Qy 1568 TCAAAATGCTGTGCTGCTACAAATTTCTATGGGCACTGTTTCGAAAACCATCTTTTAAACTCAC 1627  
Db 1592 TCAAGTGTGCTGCATACAAATTTCCCTTGGCACTCTTGTGAGAGCATCTTTTAACTCTC 1651  
Qy 1628 CAGGCCCCCTTCCCTTTCAATCCAAAGCAACATCTCTCTTATGCGACCATTTGGGCTCTGTC 1687  
Db 1652 CAGGCCCCCTTCCCTTTCAATCCAAAGCAACATCTCTCTTATGCAACAAATTTGGTGTGTC 1711  
Qy 1688 TCCCTTCAATTTGTTTCTCATTTGTTGATCTGCCAATAACAAAAAGCAATTTAGGT 1747  
Db 1712 TCCTCTTCAATTTGTTTAACTGCTTAACTTGTACAAAGTACAAAAAGCAATTTAGGT 1771  
Qy 1748 ACAGAGTCACTGTCAGATGATCCAGGTGACCTGGCCCTTGATACAGAGTACTTCTACG 1807  
Db 1772 ATGAAAGCCAGCTACAGATGATGTCAGGTGACCGCTCTCAGATAATGATCTTCTACG 1831  
Qy 1808 TTGACTTCAGGACTATGAATATGACCTTAAAGTGGAGTTCCTCGAGAGAGAACTTAGAGT 1867  
Db 1832 TTGATTTCAAGAAATATGAATATGATCTCAATGGAGTTCCTCAAGAGAAAAATTTAGAGT 1891  
Qy 1868 TTGGAAAGTCTGGGTCTGGGCTTTCGGAGGGTGAATGAACGCCACGGCTATGGCA 1927  
Db 1892 TTGGAAAGTACTAGGATCAGGTGCTTTTGGAAAAAGTGAATGAACGCAACAGCTTATGGA 1951  
Qy 1928 TTAGTAAACCGGAGTCTCAATTCAGGTGGCGTGAAGATGCTTAAAGAGAAAGCTGACA 1987  
Db 1952 TTAGCAAAACAGGAGTCTCAATTCAGGTTCCTGCTCAAAATGCTGAAAGAAAAAGCAGACA 2011  
Qy 1988 GCTGTAAAAAGAAAGCTCTCATGTCGAGCTCAAAATGATGACCCACCTGGGACACCATG 2047  
Db 2012 GCTGTAAAGAGAGCACTCATGTACAGAACTCAAGTATGATGACCCAGCTGGAGGCCAG 2071  
Qy 2048 ACAACATCGTAATCTGCTGGGGCATGCACTGTCAAGGCGCAGTGTACTTGAATTTTG 2107  
Db 2072 AGAATATTTGAAACCTGCTGGGGCGTGCACACTGTCAAGCACTTACTTGAATTTTG 2131  
Qy 2108 AATATTTGTTGTTGTCGCTCCTCACTACCTAAGAGTAAAGAGAGAGTTCACA 2167  
Db 2132 AATCTGTTGTTGTTGTTGTTCTTCACTATCTAAGAGTAAAGAGAGAAAAATTTTCA 2191  
Qy 2168 GGACATGGACAGAGATTTTAAAGAAATAATTTTCAAGTCTTACCTCTTACCTCTCCAGGCAC 2227  
Db 2192 GGACTTGGACAGAGATTTTCAAGGAAACAAATTTTCAAGTCTTACCTCTTCCATCTAC 2251  
Qy 2228 ATTCAAATTCAGCATGCTGTTTCAAGAGATTCAGTTACACCCGCCCTTTGGATCAGC 2287  
Db 2252 ATCCAAATTCAGCATGCTGTTTCAAGAGAGTTCAGATACACCCGGACTCGGATCAA 2311  
Qy 2288 TCTCAGGGTCAATGGGAAATTCATTTCTTGAAGATGAGATTGAATATGAAACCCAGA 2347  
Db 2312 TCTCAGGGCTTTCATGGGAAATTCATTTCTCTGAAAGTGAATGAAATATGAAACCCAAA 2371





TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 58...3039  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 139..3036  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 58..138  
us-08-252-517-3

Query Match 64.5%; Score 2227.2; DB 1; Length 3501;  
Best Local Similarity 79.6%; Pred. No. 0;  
Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;

QY 8 GGCTACCGCGCTCCGAGGCGCATCGCGGCGTTGGCGAGCGAGCGACCGCGCGCTGC 67  
DB 35 GGGGACCCCGGCTCCGAGGCGCATCGCGGCGTTGGC---GCGCGACGCGGCGACCGTGC 91  
QY 68 TGCTGCTTGTGTTTGTGTCAGTAATCTTGACACCGTTACAAACCAAGACCTGCCTG 127  
DB 92 CGCTGCTCGTTGTTTTTCTGCAATGATATTGGGACTATTACAAATCAAGATCTGCCTG 151  
QY 128 TGATCAAGTGTGTTTTAATCAGTCATGAGAACAAATGGCTCATCGCGGGAAGCCATCAT 187  
DB 152 TGATCAAGTGTGTTTTAATCAATCATAGAACAAATGATTCATCAGTGGGGAAGTCATCAT 211  
QY 188 CGTACCGAATGGTGGAGATCCCGCAGAAAGCTTCCAGTACCCCGAGGCGCCAGAGTG 247  
DB 212 CATATCCCATGGTATCAGAACTCCCGGAAGACCTCGGGTGTGGTGTGAGACCCAGAGCT 271  
QY 248 AAGGACGGTATATGAAGCGCCACCGTGGAGGTGGCGAGTCTGGTTCATCAACCTGC 307  
DB 272 CAGGACAGTGTAGAAAGCTCCGCTGTGGAAGTGGATGTATCTGCTTCCATCACTATGC 331  
QY 308 AAGTGCAGCTCGCACCCCGGAGACCTTCTGCTCTGGGTCTTTAAGCACAGCTCCC 367  
DB 332 AAGTGTGCTGATGCCCGCAGGAACATTTCTGCTCTGGGTCTTTAAGCACAGCTCCC 391  
QY 368 TGGGCTGCCAGCGCATTTGATTTACAAAACAGAGGAATCGTTTCCATGGCCATCTTGA 427  
DB 392 TGAATTGCCAGCCACATTTGATTTACAAAACAGAGGAGTTGTTTCCATGTGCTATTTGA 451  
QY 428 ACGTGACAGACCCAGGCGAGGAAATACCTACTTCCATATTCAGAGCGAACCGGCAACT 487  
DB 452 AAATGACAGAAACCCAAAGCTGGAGAAATACCTACTTTTATTCAGAGTGAAGCTACCAAT 511  
QY 488 ACACAGTACTCTTCACAGTGAATGTAAGAGATACACAGCTGTATGTCTAAGGAGACCTT 547  
DB 512 ACACAAATTTGTTACAGTGAATGTAAGAAATACCTGCTTTACATTAAGAGACCTT 571  
QY 548 ACTTTAGAAAGATGAAACACAGGATGCACTGCTCTGCATCTCCGAGGCTGTCCGAGC 607  
DB 572 ACTTTAGAAATGMAAACACAGAGCGCTGCTGATATCTGAGAGCGTTCCAGAGC 631  
QY 608 CCACTGGAGTGGTGTCTGAGCTCCACAGGAAAGCTGTAAAGAAAGAGCCCTG 667  
DB 632 CGATCGTGGAAATGGGTGCTTTGGCGAATTCACAGGGGGAAGCTGTAAAGAAAGTCCAG 691  
QY 668 CTGTTGTGAGAAAGGGAAGTACTTCAATGATGTTTTCGGAACAGACATCATAGTCT 727  
DB 692 CTGTTGTTAAAGAGGAAAGTGTCTTCAATGAAATTTATTTGGGACGAGACATAAGGTGCT 751  
QY 728 GTGCTAGAAATGCACTGGGCGCGGAATGCACTCAAGCTGTTTCAACATAGATCTTAAACAGG 787  
DB 752 GTGCCAAGATGAACTGGGCGAGGAAATGCACAGGCTGTTTCAATAGATCTTAAATCAA 811  
QY 788 CTCCTCAGACACACTGCCCGCAGTTATTTCTGTAAAGTGGGGAAACCTTTCTGGATCAGGT 847

DB 812 CTCCTCAGACACATGTCACAAATTTATTTCTTAAGTAGGGGAAACCTTTATGATGAAGT 871  
QY 848 GTAAGGCCATCCATGTGAACCATGGATTGGGCTCACTCTGGGAGCTGGGAAGCAAAAGCCC 907  
DB 872 GCMAAGCTGTTTCATGTGAACCATGATTCGGGCTCACCTGGGAAATTAGAAAACAAAGCAC 931  
QY 908 TGGAGGAGGCGAGCTACTTTTGAGATGAGTACTTCTCAAAACAGGACCATGATTCGGA 967  
DB 932 TCGAGGAGGCGCAACTACTTTTGAGATGAGTACTTATTTCAAAACAGGAACTATGATACGA 991  
QY 968 TTCTCTGGGCTTTGTTGTTCTTCGTTGGGAAGCAACGACGCGATATTTACACTGCTCTT 1027  
DB 992 TTCTGTTTGTGTTTGTATCATCAGTGGCAAGAAACGACACCGGATACTACATTTGTTCT 1051  
QY 1028 CCTCAAAGCACCCAGCGAGTCAAGCTTGGTGACCATCTCTAGAAAAGGGTTTATAAAGC 1087  
DB 1052 CTTCAAAGCATCCCGAGTCAATCAGCTTTGGTTACCATGCTAGGAAAGGATTTATAATG 1111  
QY 1088 CTACAGCTCGCAAGAAAGATGAAATGACCCCTGACGAAAGTTCTGCTTCTCAGTCA 1147  
DB 1112 CTACCAATTTCAAGTGAAGATTTATGAAATTTGACCAATATGAAAGATTTGTTTTCTGTCA 1171  
QY 1148 GGTTTAAAGCGTACCCACGAATCCGATGCACTGATCTTCTCTCAAGCCTCATTTCTCT 1207  
DB 1172 GGTTTAAAGCGTACCCACGAATCAGATGTACGTGGACCTTCTCTCGAAAATCATTTCTCT 1231  
QY 1208 GTGAAACAGAGAGGCTGGAGGATGGGTACAGCATATCTAAATTTTTCGATCATATAAGCA 1267  
DB 1232 GTGACCAAGGGTCTTGATTAACGGATACAGCATATCCAAATTTTGCAATATAAGCAC 1291  
QY 1268 AGCCAGGAGATACATATTTCTATGCAAAATGATGACGCCAGTTCACCAAAATGTTCA 1327  
DB 1292 AGCCAGGAGATATATATTTCCATGCAAAAATGATGATGCCCAATTTACCAAAATGTTCA 1351  
QY 1328 CGCTGAATATAAGAAAGAAACCTCAAGTGTCTGCAAAATGCTCAGCAGCCAGGCTCT 1387  
DB 1352 CGCTGAATATAAGAAAGAAACCTCAAGTGTCTGCAAAATGCTCAGCAGCCAGGCTCT 1411  
QY 1388 GTTCTCTGATGGCTACCCGCTACCTCTGCGACTGGAAGAGTGTTCGGAACAATCTC 1447  
DB 1412 GTTCTCGATGGATACCCATTCATCTTGACCTGGAAGAGTGTTCAGACAAGTCTC 1471  
QY 1448 CCAATTCGAGGAGAAATCCAGAGGAGTTTGGATATAAAAGGCTAACAGAAAGTGT 1507  
DB 1472 CCAATTCGAGAGAGATACAGAGGAGTCTGGAATAGAAAGCTAACAGAAAGTGT 1531  
QY 1508 TTGCGAGTGGGTGTGAGCAGTACTCTAAATATGATGAGCGCGGGAAGGCTTCTGG 1567  
DB 1532 TTGACACAGTGGGTGTGAGCAGTACTCTAAACATGATGAGTGAAGCCATAAAAGGCTTCTGG 1591  
QY 1568 TCRAATGCTGTGCTTACAAATTTCTATGGCAGCTCTTGGCAAAACCATCTTTTAAACTCAC 1627  
DB 1592 TCAAGTGTGTGCAATACAAATTTCCCTTGGCACATCTTGTGAGACGATCTTTTAAACTCTC 1651  
QY 1628 CAGGCCCTTCCCTTTTCATCCAGACAAACATCTCTCTATGCGACCATTTGGGCTCTGTC 1687  
DB 1652 CAGGCCCTTCCCTTTTCATCCAGACAAACATCTCTCTATGCAACAAATTTGGTGTGTC 1711  
QY 1688 TCCCTTCAATGTTGTTCTCATTTGTTGATCTGCCACAAATACAAAAGCAATTTAGT 1747  
DB 1712 TCCTCTCAATGTTGTTTAAACCTGCTAAATTTGTCACAAGTACAAAAGCAATTTAGT 1771  
QY 1748 ACAGAGTCAGCTCAGATGATCCAGGTACTGCGCCCTCGGATTAACAGTACTTCTACG 1807  
DB 1772 ATGAAGCCAGCTACAGATGATCAGGTGACCGGCTCCTCAGATAATGAGTACTTCTACG 1831  
QY 1808 TTGACTTTCAGGACTATGAATATGACCTTAAGTGGGATTTCCCGAGAGAGAACTTAGAGT 1867  
DB 1832 TTGATTTACAGAAATATGAATATGATCTCAATGGGATTTTCAAGAGAAATTTTAGAGT 1891  
QY 1868 TTGGGAAGGCTCTGGGCTCTGGCGCTTTTCGGGAGGGTGAATGAAACCCACGCGCTATGCA 1927

1892 TTGGAGAGGTACTAGGATCAGTGCTTTTGGAAAGTGATGAACGCAACAGCTTATGGAA 1951  
1928 TTAGTAAACGGGAGTCTCAATTCAGGTGGCGTGAAGTGTCTAAAGAGAAAGCTGACA 1987  
1952 TTAGCAAAACAGGAGTCTCAATTCAGGTGGCGTCAAAATGCTGAAGAGAAAGCAGACA 2011  
1988 GCTGTGAAAAAGAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGACACCATG 2047  
2012 GCTGTGAAAGAGAGCACTCATGTGAGAACTCAAGATGATGACCCAGCTGGGAGCCACG 2071  
2048 ACAACATCGTAATCTGTGGGGGATGACACTGTCTAGGGCCAGTGTACTGTGATTTTGG 2107  
2072 AGAATATTTGTGAACCTGTGGGGGCTGACACTGTCTAGGACCAATTTACTTGATTTTG 2131  
2108 AATATTGTTGTATGTTGACCTCTCACTACTAATCTTAAGAAACATAATTTTCACTTCTT 2167  
2132 AATACTGTTGTATGTTGATCTTCTCACTACTAATCTTAAGAAAGTAAAGAGAGAAATTT 2191  
2168 GGACATGGACAGAGATTTTAAAGAAACATAATTTTCACTTCTTCACTTCTTCCAGGCAC 2227  
2192 GGACTTGGACAGAGATTTTCAAGGACACAAATTTTCACTTCTTCCCACTTTCCAACTAC 2251  
2228 ATTCAAAATTCAGCATGCTGTTTCAAGAGAGTTTCAAGAGAGTTTCAAGAGAGTTTCA 2287  
2252 ATCCAAATTCAGCATGCTGTTTCAAGAGAGTTTCAAGAGAGTTTCAAGAGAGTTTCA 2311  
2288 TCTCAGGGTCAATGGGAATTCATTTTCACTCTGAAGATGAAATTTGAATTTGAAGATTT 2347  
2312 TCTCAGGGCTTCAATGGGAATTCATTTTCACTCTGAAGATGAAATTTGAATTTGAAGATTT 2371  
2348 AGAGGCTGGCAG 2407  
2372 AAGGGCTG-----GAAG 2425  
2408 GCTTTGCGTACCAAGTGGCCAAAGGCAATGGAATTTCTTGGAGTTCAAAGTGTGTCCACA 2467  
2426 GCTTTGCAATCAAGTTGGCCAAAGGCAATGGAATTTCTTGGAGTTCAAAGTGTGTCCACA 2485  
2468 GAGACCTGGCAGCAGGAG 2527  
2486 GAGACCTGGCAGCAGGAG 2545  
2528 TTGGACTGGCCCGCAGAGACATCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2587  
2546 TTGGATTGGCTCAGATATATCAGATATATCAGATATATCAGATATATCAGATATATCAG 2597  
2588 TGCCCGTGAAGTGGATGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2605  
2606 TGCCCTGTAATAATGGATGGCCCGCCGAAAGCCTGTTGAGGCATCTACACCAATCAAGAGTG 2647  
2648 AGCTGTGGTCTTACGGCATCTTCTCTGGAGATATTTTCACTGGGTGTGAACCTTACC 2707  
2666 ATGCTGTGGTCATATGGAATATTTACTTGGGAAATCTTCTCACTTGGTGTGAATCTTACC 2725  
2708 CTGGCAATCTCTGTCAGCGTAACTTCTATAACTGATTCAGAGTGAATTTAAATGGAGC 2767  
2726 CTGGCAATCCGGTGTATGCTAACTTCTCAAACTGATTCAAAATGGAATTTAAATGGATC 2785  
2768 AGCCATTTCTATGCCAGAGAGGATATATCTTTGTAATGCAATCTCTGCGGCTTTTGGACT 2827  
2786 AGCCATTTTATGCTACAGAGAAATATATCAATTAATGCAATCTCTGCGGCTTTTGGACT 2845  
2828 CAAAGAGCGGCGCTCTTCCCAACCTGACTTCAATTTTAAAGATGTCTAGCTGGCAGAGG 2887  
2846 CAAAGAGAGCGGCGCTCTTCCCAACCTGACTTCAATTTTAAAGATGTCTAGCTGGCAGAGG 2905  
2888 CAGAAGAG-----CATGTATCAGAAATGGATGGCGGTGTTTGGGAATGTTCTCTCACCTACC 2924  
2906 CAGAAGAGCGATGTATCAGAAATGGATGGCGGTGTTTGGGAATGTTCTCTCACCTACC 2965  
2925 AAAACAGGCGGCGCTCTTCCCAACCTGACTTCAATTTTAAAGATGTCTAGCTGGCAGAGG 2983  
2966 AAAACAGGCGGCTCTTCCCAACCTGACTTCAATTTTAAAGATGTCTCTCCGAGGCTCAGG 3025

QY 2984 T-GAAGATTCTACAGAGAAAGATTAGCGAGGAGCCTTGGACCCCGCC-----ACCTAG 3038  
Db 3026 TCGAAGATTCTAGAGGAAATTTAGTTTAAAGGACTTCATCCCTCCACCTATCCCTAA 3085  
QY 3039 CAGGCTGTAGACCGCAGAGCAAGATTAGCCTCGCCTCT--GAGGAAGCGCCTACAGCG 3096  
Db 3086 CAGGCTGTAGATTACCAAAACAGATTAAATTTTCATCACTAAAGAAATCTATTATCAAC 3145  
QY 3097 CGTTGCTTCGCTGGACTTTTCTCTAGATGCTGTCTGCCATTACTC-----CAAAGTGA 3149  
Db 3146 TGCTGCTTCCAGACTTTTCTCTAGAGCCGCTCTGCGTTTACTCTTGTGTTTCAAAGGGA 3205  
QY 3150 CTTCTATAAAATCAAACTCTCTCGCACAGCGCGGAGAGCAATAATGAGACTTGTGG 3209  
Db 3206 CTTTGTAAATCAAAATCATCTCTGTCACAGCGCAGGAGGAGCTGATATGAACTTTATTG 3265  
QY 3210 TGAGCCCGCTACCTCGGGGCTTTTCCACGAGCTTGAGGGGAAAGCCATGTATCTGAAA 3269  
Db 3266 GAGCAATTGATCTGCATCCAAAGGCTTCTCAGCGCGGCTTGAGTGAATTTGTCTACCTGAAG 3325  
QY 3270 TATAGTATATTCTTGTAATAACGTAACCAAAACCCGCTTTTTCGCTAAGGAAAGC 3329  
Db 3326 TACAGTATATTCTTGTAATAACGTAACCAAAACCCGCTTTTTCGCTAAGGAAAGC 3378  
QY 3330 TAAATATGATTTTAAATACTATGTTTAAATACTATGTAACTTTTTCATCTATTATAG 3389  
Db 3379 TAATATGATTTT--AAGTCTATGTTTAAATACTATGTAACTTTTTCAGCTATTATAG 3436  
QY 3390 TGATATATTATGGAATGGAATAAACTTTCTACTGTAAATAAAATAAAATAAAATAAA 3449  
Db 3437 TGATATATTATGAGTGGGATATAAATTTCTACTACAGAAATAAATAAATAAATAA 3496  
QY 3450 AAAA 3453  
Db 3497 AAAA 3500

Search completed: August 28, 2003, 05:51:50  
Job time : 296.55 sec

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2003, 07:20:58 ; Search time 32.9834 Seconds  
(without alignments)  
2892.346 Million cell updates/sec

Title: US-09-919-408A-2

Perfect score: 5264

Sequence: 1 MRALAQRSDRLLLLVLSV.....RGLRAQSPORQVQKIHRS 992

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 76.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5264	100.0	992	2 A39931	protein-tyrosine k
2	5102	96.9	1000	2 S18827	Flt3 protein - mou
3	4408.5	83.7	993	2 A36873	protein-tyrosine k
4	4258.5	23.9	980	1 TVCTWD	macrophage colony-
5	1247	23.7	941	1 TVMWND	protein-tyrosine k
6	1232.5	23.4	978	1 A49814	protein-tyrosine k
7	1223	23.2	975	1 TVMSKT	protein-tyrosine k
8	1221.5	23.2	972	1 TVHUND	macrophage colony-
9	1220	23.2	954	2 T51703	c-kit-related kina
10	1215	23.1	977	2 T45877	protein-tyrosine k
11	1210	23.0	978	2 S16385	macrophage colony-
12	1198.5	22.8	976	1 TVMSND	macrophage colony-
13	1196	22.7	976	1 TVHUKT	protein-tyrosine k
14	1188.5	22.6	975	2 T30816	macrophage colony-
15	1181	22.4	960	1 JN0677	protein-tyrosine k
16	1157.5	22.0	1088	1 PFRTGA	platelet-derived g
17	1152.5	21.9	1089	1 PFHUGA	platelet-derived g
18	1132	21.5	1087	2 T51552	platelet-derived g
19	1123.5	21.3	1089	1 S33727	platelet-derived g
20	1098	20.9	1098	1 PPMSEB	platelet-derived g
21	1078	20.5	1106	1 PFHUGB	platelet-derived g
22	1060	20.1	1048	2 T30815	platelet-derived g
23	998.5	19.0	1338	2 S09982	protein-tyrosine k
24	984.5	18.7	1336	2 T60598	Fit-1 tyrosine kin
25	983.5	18.7	1333	2 T78875	receptor tyrosine
26	980	18.6	1379	2 JC4954	vascular endotheli
27	970	18.4	1330	2 S49010	embryonic receptor
28	967.5	18.4	790	1 FOMVHZ	gag-kit polyprotei
29	965	18.3	1348	2 S51656	vascular endotheli

30	954.5	18.1	1298	2 A48999	protein-tyrosine k
31	954.5	18.1	1356	2 JC1402	protein-tyrosine k
32	952	18.1	1363	2 I58375	protein-tyrosine k
33	926.5	17.6	1367	2 A41228	protein-tyrosine k
34	822	15.6	160	2 A39061	protein-tyrosine k
35	815	15.5	813	1 A49123	fibroblast growth
36	810.5	15.4	823	2 B35963	protein-tyrosine k
37	798	15.2	797	2 S38579	fibroblast growth
38	791	15.0	821	1 TVHUF2	fibroblast growth
39	791	15.0	824	2 S24108	protein-tyrosine k
40	789	15.0	806	2 A35963	protein-tyrosine k
41	788.5	15.0	769	2 S16236	fibroblast growth
42	788.5	15.0	822	2 A45081	fibroblast growth
43	780.5	14.8	822	2 A41794	keratinocyte growt
44	779.5	14.8	457	2 S44269	platelet-derived g
45	777	14.8	821	1 TVMSBK	fibroblast growth

#### ALIGNMENTS

##### RESULT 1

A39931

protein-tyrosine kinase (EC 2.7.1.112) flk-2, hematopoietic - mouse

C;Species: Mus musculus (house mouse)

C;Date: 17-Jan-1992 #sequence\_revision 17-Jan-1992 #text\_change 04-Feb-2000

C;Accession: A39931

R;Matthews, W.; Jordan, C.T.; Wiegand, G.W.; Pardoll, D.; Lemischka, I.R.

Cell 65, 1143-1152, 1991

A;Title: A receptor tyrosine kinase specific to hematopoietic stem and progenitor cell-

A;Reference number: A39931; MUID:91292518; PMID:1648448

A;Accession: A39931

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-992 <MAT>

A;Cross-references: GB:M64689; NID:G193327; PIDN:AAA37634.1; PID:G193328

C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol

C;Keywords: ATP; phosphotransferase; transmembrane protein; tyrosine-specific protein k

F;609-953/Domain: protein kinase homology <KIN>

F;617-625/Region: protein kinase ATP-binding motif

Query Match 100.0%; Score 5264; DB 2; Length 992;  
Best Local Similarity 100.0%; Pred. No. 8.6e-263;  
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MRALAQRSDRLLLLVLSVMIETVTNQLPVTKVLISHENNGSSAGKPSSYRVMVRGS	60
Db	1	MRALAQRSDRLLLLVLSVMIETVTNQLPVTKVLISHENNGSSAGKPSSYRVMVRGS	60
Qy	61	PEDLOCTPRQSEGTVEAATVEAEGSITLQVQLATPGDLSCLWVFKHSSLCQCPHFD	120
Db	61	PEDLOCTPRQSEGTVEAATVEAEGSITLQVQLATPGDLSCLWVFKHSSLCQCPHFD	120
Qy	121	LQNRGVSMALNVTTQAGEYLLHTQSERANTVTLTVANVTQTLVLRPYRKMENQ	180
Db	121	LQNRGVSMALNVTTQAGEYLLHTQSERANTVTLTVANVTQTLVLRPYRKMENQ	180
Qy	181	DALLCISEGVPEPTVEWLCSHRESCKEGPAVRKEEKLHFGTDRCCARNALGR	240
Db	181	DALLCISEGVPEPTVEWLCSHRESCKEGPAVRKEEKLHFGTDRCCARNALGR	240
Qy	241	ECTKLFITDLNQAPQSTLPOLFLKVGEPLEIRCKAIHVNHGFGLTWELEKALEEGSYFE	300
Db	241	ECTKLFITDLNQAPQSTLPOLFLKVGEPLEIRCKAIHVNHGFGLTWELEKALEEGSYFE	300
Qy	301	MSYTSNTRTMIRILLAPVSSVGRNDTGYTCSSKHPQSALVTILEKGFINATSSQBEY	360
Db	301	MSYTSNTRTMIRILLAPVSSVGRNDTGYTCSSKHPQSALVTILEKGFINATSSQBEY	360
Qy	361	EIDPYEKFCPSVRKAYPRIRCTWIFSQAQFPCEQGLDGYSISKFCDHKNKGEYIFY	420
Db	361	EIDPYEKFCPSVRKAYPRIRCTWIFSQAQFPCEQGLDGYSISKFCDHKNKGEYIFY	420

```
QY 421 AENDDAQFTKMTLNIRKKPOVLANASQASCSGSDGYPLPSWTWKKCDKSPNCTEIP 480
DB 421 AENDDAQFTKMTLNIRKKPOVLANASQASCSGSDGYPLPSWTWKKCDKSPNCTEIP 480
QY 481 EGVNKKANRVKFGQWSSSTLNSEAGKGLLVKCCAYNSMGTSCTETIFLNSPGPPFIQ 540
DB 481 EGVNKKANRVKFGQWSSSTLNSEAGKGLLVKCCAYNSMGTSCTETIFLNSPGPPFIQ 540
QY 541 DNISFYATIGLCLPFIIVLIVLI CHYKQKQRYESQLOMI QVTGPDLDNEYFVDPDRDYEY 600
DB 541 DNISFYATIGLCLPFIIVLIVLI CHYKQKQRYESQLOMI QVTGPDLDNEYFVDPDRDYEY 600
QY 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISKTVGSIQVAVKMLKEKADSCKEALM 660
DB 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISKTVGSIQVAVKMLKEKADSCKEALM 660
QY 661 SELKXMTLGHHDNIVNLGACTLSGPPVLI FEYCCYGDLNLYRSKKEKPHRTWTEIFK 720
DB 661 SELKXMTLGHHDNIVNLGACTLSGPPVLI FEYCCYGDLNLYRSKKEKPHRTWTEIFK 720
QY 721 EHNFSYPTFOAHSNSMPGSGREVQLHPDLQSGFNGNSIHSDEIEYENQKRLAESEE 780
DB 721 EHNFSYPTFOAHSNSMPGSGREVQLHPDLQSGFNGNSIHSDEIEYENQKRLAESEE 780
QY 781 EDNLVTFEDLLCFAYQVAKGMEFLEFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840
DB 781 EDNLVTFEDLLCFAYQVAKGMEFLEFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840
QY 841 SDSSYVVRGNARLPVKWMAPELSEFEGYTFVWQSCWAFDSKRPSPFNLTSLFLGQLAEEACIRT 900
DB 841 SDSSYVVRGNARLPVKWMAPELSEFEGYTFVWQSCWAFDSKRPSPFNLTSLFLGQLAEEACIRT 900
QY 901 FYKLIQSGFQWMEQPFYATGTYFVWQSCWAFDSKRPSPFNLTSLFLGQLAEEACIRT 960
DB 901 FYKLIQSGFQWMEQPFYATGTYFVWQSCWAFDSKRPSPFNLTSLFLGQLAEEACIRT 960
QY 961 SIHLPKQAAPQOORGLRAQSPORQVKIHRERS 992
DB 961 SIHLPKQAAPQOORGLRAQSPORQVKIHRERS 992

RESULT 2
SI8827
FIt3 protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 24-Sep-1999
C;Accession: SI8827
R;Rosnet, O.; Marchetto, S.; deLapeyriere, O.; Birnbaum, D.
Oncogene 6, 1641-1650, 1991
A;Title: Murine FIt3, a gene encoding a novel tyrosine kinase receptor of the PDGFR/CSF1
A;Reference number: SI8827; MUID:92019834; PMID:1656368
A;Accession: SI8827
A;Molecule type: mRNA
A;Residues: 1-1000 <ROS>
A;Cross-references: EMBL:X59398; NID:950978; PIDN:CAA42041.1; PID:g50979
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C;Keywords: Atp
F;609-953/Domain: protein kinase homology <KIN>
F;617-625/Region: protein kinase Atp-binding motif

Query Match 96.9%; Score 5102; DB 2; Length 1000;
Best Local Similarity 96.7%; Pred. No. 1.8e-254;
Matches 968; Conservative 6; Mismatches 17; Indels 10; Gaps 2;

QY 1 MRALAQRSDRLLLLVLSVMILETVTNQDLIPVKLVLSHNNNGSSAGKPSRYMRGRS 60
DB 1 MRALAQRSDRLLLLVLSVMILETVTNQDLIPVKLVLSHNNNGSSAGKPSRYMRGRS 60
QY 61 PEDLOCTPRROSEGTVEAAATVEAEGSITLQVLAATPGDLSCLVWPKHSLGCOQPHD 120
DB 61 PEDLOCTPRROSEGTVEAAATVEAEGSITLQVLAATPGDLSCLVWPKHSLGCOQPHD 120
QY 121 LQNRGIVSMALNVTETQAGEYLLHIQSEANVTYVLFVNVRDITQLYVLRPRYFKRMENQ 180
```

```
DB 121 LQNRGIVSMALNVTETQAGEYLLHIQSEANVTYVLFVNVRDITQLYVLRPRYFKRMENQ 180
QY 181 DALLCISGVPPEPTVEVWVLCSSHRESCKEKGPAVVRKEEVLHFGTDIRCCARNALGR 240
DB 181 DALLCISGVPPEPTVEVWVLCSSHRESCKEKGPAVVRKEEVLHFGTDIRCCARNALGR 240
QY 241 ECTKLFTIDLNAQPSQSTLPQLFLKVGKPLWIRCKAIHVNHGFGLTWELEDKALESGSYFE 300
DB 241 ECTKLFTIDLNAQPSQSTLPQLFLKVGKPLWIRCKAIHVNHGFGLTWELEDKALESGSYFE 300
QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTTCSSKHPQSALVTILEKGFNATSSQBEY 360
DB 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTTCSSKHPQSALVTILEKGFNATSSQBEY 360
QY 361 EIDPEYKPCSVRKPAYPRIRCTWIFSOASPPCQRLGEGYSISKPCDHKNKPGEYIFY 420
DB 361 EIDPEYKPCSVRKPAYPRIRCTWIFSOASPPCQRLGEGYSISKPCDHKNKPGEYIFY 420
QY 421 AENDDAQFTKMTLNIRKKPOVLANASQASCSGSDGYPLPSWTWKKCDKSPNCTEIP 480
DB 421 AENDDAQFTKMTLNIRKKPOVLANASQASCSGSDGYPLPSWTWKKCDKSPNCTEIP 480
QY 481 EGVNKKANRVKFGQWSSSTLNSEAGKGLLVKCCAYNSMGTSCTETIFLNSPGPPFIQ 540
DB 481 EGVNKKANRVKFGQWSSSTLNSEAGKGLLVKCCAYNSMGTSCTETIFLNSPGPPFIQ 540
QY 541 DNISFYATIGLCLPFIIVLIVLI CHYKQKQRYESQLOMI QVTGPDLDNEYFVDPDRDYEY 600
DB 541 DNISFYATIGLCLPFIIVLIVLI CHYKQKQRYESQLOMI QVTGPDLDNEYFVDPDRDYEY 600
QY 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISKTVGSIQVAVKMLKEKADSCKEALM 660
DB 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISKTVGSIQVAVKMLKEKADSCKEALM 660
QY 661 SELKXMTLGHHDNIVNLGACTLSGPPVLI FEYCCYGDLNLYRSKKEKPHRTWTEIFK 720
DB 661 SELKXMTLGHHDNIVNLGACTLSGPPVLI FEYCCYGDLNLYRSKKEKPHRTWTEIFK 720
QY 721 EHNFSYPTFOAHSNSMPGSGREVQLHPDLQSGFNGNSIHSDEIEYENQKRLAESEE 780
DB 721 EHNFSYPTFOAHSNSMPGSGREVQLHPDLQSGFNGNSIHSDEIEYENQKRLAESEE 780
QY 781 EDNLVTFEDLLCFAYQVAKGMEFLEFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840
DB 781 EDNLVTFEDLLCFAYQVAKGMEFLEFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840
QY 841 SDSSYVVRGNARLPVKWMAPELSEFEGYTFVWQSCWAFDSKRPSPFNLTSLFLGQLAEEACIRT 900
DB 841 SDSSYVVRGNARLPVKWMAPELSEFEGYTFVWQSCWAFDSKRPSPFNLTSLFLGQLAEEACIRT 900
QY 901 FYKLIQSGFQWMEQPFYATGTYFVWQSCWAFDSKRPSPFNLTSLFLGQLAEEACIRT 960
DB 901 FYKLIQSGFQWMEQPFYATGTYFVWQSCWAFDSKRPSPFNLTSLFLGQLAEEACIRT 960
QY 961 -----SIHLPKQAAPQOORGLRAQSPORQVKIHRERS 992
DB 961 MGNVPEHPHSYQNRRLPSREAGS-BPPSPQAQVKIHRERS 1000

RESULT 3
A36873
protein-tyrosine kinase (EC 2.7.1.112) STK-1 precursor - human
N;Alternate names: stem cell tyrosine kinase 1
C;Species: Homo sapiens (man)
C;Date: 03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change 04-Feb-2000
C;Accession: A36873
R;Small, D.; Levenstein, M.; Kim, E.; Carow, C.; Amin, S.; Rockwell, P.; Witte, L.; Bur
Proc. Natl. Acad. Sci. U.S.A. 91, 459-463, 1994
A;Title: STK-1, the human homolog of Flk-2/Flt-3, is selectively expressed in CD34(+) h
A;Reference number: A36873; MUID:94119906; PMID:7507245
A;Accession: A36873
A;Status: preliminary
```

A:Molecule type: mRNA  
A:Residues: 1-993 <SMA>  
A:Cross-references: GB:002687  
A:Note: in the authors translation, an additional residue Ala is shown after 420-Ala and  
C:Genetics:  
A:Map position: 13q12  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
C:Keywords: ATP; glycoprotein; phosphotransferase; transmembrane protein; tyrosine-speci  
F:608-950/Domain: protein kinase homology <KIN>  
F:616-624/Region: protein kinase ATP-binding motif

Query Match 83.7%; Score 4408.5; DB 2; Length 993;  
Best Local Similarity 84.0%; Pred. No. 6.8e-219;  
Matches 836; Conservative 57; Mismatches 89; Indels 13; Gaps 6;

QY 1 MRALAQSRDRLLILVLSVMILETVTNQDLVPIKCVLISHENNGSAGKPSVYRMVGRS 60  
DB 1 MPALA-RDAGTVPLLVVFSAMIFGTIINQDLVPIKCVLINHKNDSSVGKSSYPWVSES 59

QY 61 PEDLOCTPRQSEGTYYEATATVEVAESGSLTQVOLATPGDLSCLWFKHSSLCQPHD 120  
DB 60 PEDLGALRQSSGTYVEAAAEVDVSGASITLQVLDAPGNISCLWFKHSSLCQPHD 119

QY 121 LQNRGVSMALNVTTQAGEYLLHIOSEANVTYLTAVNRDTQVLYLREPYPKRMENQ 180  
DB 120 LQNRGVSMVILKMTQAGEYLLFIOSEATNYTLFTVIRNTLLYLRPRYPKRMENQ 179

QY 181 DALLCTSEGVPETVEWVLCSSHRESCKEPAVVRKEEVKHLFGTDIRCCARNALGR 240  
DB 180 DALVCISESPEPIVEWVLCDSQESCKERSPAVVRKEEVKHLFGTDIRCCARNELGR 239

QY 241 ECTKLTIDLNOAQPOSTLPFLKVGPELWIRCKAIHVNHFGLTWELKALBEGSYFE 300  
DB 240 ECTRLPTIDLNQPTTLPFLKVGPELWIRCKAVHNVHGFGLTWELKALBEGSYFE 299

QY 301 MSTYSTNRTMIRILLARVSSVGRNDTGYTCSSSKHPSQSALVTILEKGFNATSSOEY 360  
DB 300 MSTYSTNRTMIRILFAVSSVARNDTGYTCSSSKHPSQSALVTILEKGFNATSSBEDY 359

QY 361 EIDPYEKFCFSVRFKAYPRIRCTWIFSOAGFPFCBQGLDGYSTKFCDHKNKGEYIFY 420  
DB 360 EIDQYEEFCFSVRFKAYPRIRCTWIFSRKFPCEQKGLDNGYSISKFCNKHQGEYIFH 419

QY 421 -AENDDAQFTKMTFLNIRKPKQVLNANASASQSCSDGYPLPSWTWKCKSDKSPNCTEEI 479  
DB 420 AENDDAQFTKMTFLNIRKPKQVLNANASASQSCSDGYPL-SWTWKCKSDKSPNCTEEI 478

QY 480 PEGVNNKANRKPFGQWVSSSTLNMSEAGKLLVKKCAVNSMGTSCTIFLNSGPPFFI 539  
DB 479 TEGVNNKANRKPFGQWVSSSTLNMSEAIKGLVKKCAVNSLGTSCTIFLNSGPPFFI 538

QY 540 QDNISFYATIGLCLPFTIVLVILICHKYKKOPRYESQLOMIVTQPLDNEYFYVDFRDYE 599  
DB 539 QDNISFYATIGVCLLFIVLVLLICHKYKKOPRYESQLOMIVTQPLDNEYFYVDFREYE 598

QY 600 YDLKWEPPRENLEFGKVLGSGAFGRVNNATAYGISTGVSTQVAVKMLKERADSCKEAL 659  
DB 599 YDLKWEPPRENLEFGKVLGSGAFGRVNNATAYGISTGVSTQVAVKMLKERADSSREAL 658

QY 660 MSELKMMTHIGHNDINVLIGACTLSGPIYILPEYCCYGDLLNLYLRKREKFRHTWTIEF 719  
DB 659 MSELKMMTOLGSHENINVLIGACTLSGPIYILPEYCCYGDLLNLYLRKREKFRHTWTIEF 718

QY 720 KEHNFSSYPFOAHSSNMPGSRVOLPPDOLSGFNGNSIHSEDETEYENQKRLAEE 779  
DB 719 KEHNFSSYPFOAHSSNMPGSRVOLPPDOLSGFNGNSIHSEDETEYENQKRL--EE 776

QY 780 EEDLNVLTFEDLLCFAYQVAKMEFLFEFKSCVHRDLAARNVLVTHGKVKICDFGLARDI 839  
DB 777 EEDLNVLTFEDLLCFAYQVAKMEFLFEFKSCVHRDLAARNVLVTHGKVKICDFGLARDI 836

QY 840 LSDSSYVVRGNARLPVKWMAPESEFEGITYTIKSDVWSYGILLWEIFSLGVNYPGIPVDA 899  
DB 837 MSDSNYVVRGNARLPVKWMAPESEFEGITYTIKSDVWSYGILLWEIFSLGVNYPGIPVDA 896

QY 900 NFYKLIOSGKMSQPPFYATGIIYFMOSCWAFDSRKPSNPNTSLFGLCOLABEEAC-- 957  
DB 897 NFYKLIONGFMQPPFYATBEIYIIMOSCWAFDSRKPSNPNTSLFGLCOLADAEAWYQ 956

QY 958 -----IRTSFHLPKQAAPQORG-GLRAQSPORQVK 986  
DB 957 NVDRVSECHTYQNRRPFPREMDLGLLSPOAQVE 991

RESULT 4

TVCTMD

macrophage colony-stimulating factor 1 receptor precursor - cat

N:Contains: protein-tyrosine kinase (EC 2.7.1.12) csfir/fms

C:Species: Felis silvestris catus (domestic cat)

C:Date: 31-Dec-1989 #sequence\_revision 31-Dec-1989 #text\_change 13-Jun-1997

C:Accession: A31636

R:Woolford, J.; McAuliffe, A.; Rohrschneider, L.R.

Cell 55, 965-977, 1988

A:Title: Activation of the feline c-fms proto-oncogene: multiple alterations are requir  
A:Reference number: A31636; MUID:89077553; PMID:2849512

A:Accession: A31636

A:Molecule type: mRNA

A:Residues: 1-980 <WOO>

A:Cross-references: EMBL:X03663

C:Genetics:

A:Gene: fms

C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology  
C:Keywords: Atp; autophosphorylation; glycoprotein; kinase-related transforming protein  
fic protein kinase

F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-980/Product: macrophage colony-stimulating factor 1 receptor #status predicted <WA

F:24-980/Domain: extracellular #status predicted <EXT>

F:35-86/Domain: immunoglobulin homology <IMM1>

F:120-179/Domain: immunoglobulin homology <IMM2>

F:217-280/Domain: immunoglobulin homology <IMM3>

F:316-381/Domain: immunoglobulin homology <IMM4>

F:410-484/Domain: immunoglobulin homology <IMM5>

F:510-534/Domain: transmembrane #status predicted <TMM>

F:535-980/Domain: intracellular #status predicted <INT>

F:577-915/Domain: protein kinase homology <KIN>

F:585-593/Region: protein kinase ATP-binding motif

F:42-84,127-177,224-278,417-482/Disulfide bonds: #status predicted

F:45,73,94,153,275,335,410,477,490/Binding site: carbohydrate (Asn) (covalent) #sta

F:613,630,776/Active site: Lys, Glu, Asp #status predicted

F:781,794/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 23.9%; Score 1258.5; DB 1; Length 980;

Best Local Similarity 32.7%; Pred. No. 2.9e-57;

Matches 328; Conservative 168; Mismatches 353; Indels 153; Gaps 32;

QY 75 TVVEAAATVEAESGSLTQVOLATPGDLSCL-----W---VPKHSSIGCOPHFQDQNRG 125

DB 13 TAMHAQGVPIQSPGPELVWEPGTVTLRCVNGSVWDGPISPHNLDLPPSSI----- 68

QY 126 IVSMAILNVTTQAGEYLLHI-----QSERANYTVLTVNVRDTQLYVL--RRPY----- 173

DB 69 ---LTTNNAIFONTGTY--HCTEPNPGGNATI-----HLYVKDPARPKWLAQE 114

QY 174 FRKMENODALL-CISEGVPEPTVEWVLCSSHRESCKEGBPAVVRKEEVKHLFGTDI-- 230

DB 115 VTVLEGQDALIPCL---LTPDALE-----AGVSLVRVRGRPVLRTQNTYSFSPWHGFTTHK 166

QY 231 -----RCCARNALGRECTKL-----FTIDLNAQPOSTL--POLFLKV-GEPLWIRCK 274

DB 167 AKFTENHYOCSAR--VDGRTVTSMGIWKLQKQKISGATLTLLEPAELVRIQGEAAQIVCS 225

QY 275 AIHVNHFGLTWELKALBEGSYFEMSTYSTNRTMIRILLAFVSSVGRNDTGYTCSSS 334

DB 226 ASNIDVNFVSLRHGDTKL---TISQOSDFHDNRQ-KVLTNLNLDHVSFDQAGNYSCTAT 281

QY 335 K---HPSQSALVTILEKGFNATSSQE-EYEIDPYEKFCFSVRPKAYPRIRC-TWI----- 385

F;217-280/Domain: immunoglobulin homology <IMW3>  
F;316-381/Domain: immunoglobulin homology <IMW4>  
F;410-484/Domain: immunoglobulin homology <IMW5>  
F;510-534/Domain: transmembrane #status predicted <TMM>  
F;535-941/Domain: intracellular #status predicted <INT>  
F;577-915/Domain: protein kinase homology <KIN>  
F;585-593/Region: protein kinase ATP-binding motif  
F;42-84,127-177,224-278,417-482/Diulfide bonds: #status predicted  
F;45,73,94,153,275,286,302,335,410,477,490/Binding site: carbohydrate (Asn) (covalent)  
F;613,630,776/Active site: Lys, Glu, Asp #status predicted  
F;781,794/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 23.7%; Score 1247; DB 1; Length 941;  
Best Local Similarity 33.2%; Pred.No.1.le=56;  
Matches 325; Conservative 162; Mismatches 144; Gaps 31;

QY 75 TVYEAAATVEVAGSGSITIQVLATPGDLSCL-----W---VFKHSSLCGCPHFLQNRG 125  
DB 13 TAWHAQGVVTPQSPGELVVRPGTIVLRGVNGSVVDGFISSHWNLDLDPSSI----- 68  
QY 126 IVSMAILANVTQAGEYLLHI-----QSERANYTVLFTVNVRDQLVYL--RRPY----- 173  
DB 69 ---LTTNNAFTQNTGY--HCTEPGNRGGNATI-----HLVVKDPAHPMKVLAQE 114  
QY 174 FRKMENQDALL-CISGVPEPTVEVWLCSHRESCKEGPAAVVRKEEKVLHELFGTDI-- 230  
DB 115 VTVLEGGDALLPCL-----LTDPALE-----AGVSLVRVRGPFVLRQTNYSFSFWHGFTTHK 166  
QY 231 -----RCARNALGSECTKL-----FTIDLNAQPSQL-PQLFLKV-GEPLWIRCK 274  
DB 167 AKFIENHVYQCSAR-VDGRVTVMGIVLWKQDLSGPAATLLEPAELVRIQGEAAQIVCS 225  
QY 275 AIHVNHGFLTWELDKALEBEGSFEMSTYSTNRIMIRILLAFVSSVGRNDTGYVTCSSS 334  
DB 226 ASNIDNVDFVSLRHGDTKL---TTSQQSDFDHNRVQ-KVLINLDHVSFQDAGNVSCTAT 281  
QY 335 K---HPSQSALVTILEKGFINATSQE-EYRIDPYEKFCFVRKAYPIRC-TWI----- 385  
DB 282 NAWGNHSA-SMVRVVSAYSNLTSBSQLLEQVTGVEKVDLQVKVEAYPGLESFNWYTLGP 341  
QY 386 FSQASFPCEQRLDGYSISKFD-----HKNKPGEYIFYAENDDAQFTKMTLNIARKPQ 441  
DB 342 FSDYQDKLDFVTVIKDTYRTYTLSLPLRKRSESGRYSLARNAGQGNALTFELTLRYPPE 401  
QY 442 VLANASASQAS-----CSSDGYPLPSWTWKKCSKSPNCTEE---IPEGVNNKANRKYFG 494  
DB 402 VRVWTLLNGSDTLTLCESAGYQPQSVTWVQCRSHTRCDDESAGLVLEDSHSEVLISQVFPY 461  
QY 495 QWVSSTLNMGEAGKGLLVKCAVNSMGTSCETIFLNSPGPPFIQDNISFYATIGLC-- 552  
DB 462 EVIVHSLLAIGTLEHNRITYECRAFNVSNGSOTFWPISIGAHTPLPDELLFTPVLLTCS 521  
QY 553 -LPFIWLVILVICHYKQFYESOLQMIQVTGPLDNBYFYVDFRDVEYDLKWFPPRNL 611  
DB 522 INALLLLLLLLLXYKQPKYQVRWKIIE--SYEGNSYTFIDTQLPYNEKWFPPRNL 579  
QY 612 BFGKVLGSAFGVRVNATAYGISTGVSIOVAVKMLKEKADSCKEALMSLKMWTHLGH 671  
DB 580 QGKTLGTGAFGVVVEATFGLGKEDAVLKVAVKMLKSTAHADKEALMSLKMTHLGH 639  
QY 672 HDNIVNLLGACTLSGFPVYLI FEYCCYGDLNLYLRKRE-----KF 711  
DB 640 HENIVNLLGACTHGGFVLVITTEYCCYGDLNLFARQAEAMPGPSILSVQDPEAGAGYKNI 699  
QY 712 HRTWTEIPKEHNFYSYPTQAHNSMMPGSREVOLHPPDLQDSFGNGNSIHSEDEIYEN 771  
DB 700 HLEKKYVRDRSGFS-----SQGVDTYVEMRPVSTSSGNSDPSSEDL----- 740  
QY 772 QKRLAESEEDLNVLTFFEDILCFAYOVAKNGEFLFKSCVHRDLAARNVLTHGKVYKIC 831  
DB 741 -----GKEGRLPLELRLHFSQVAGMAFLASKNCHIRDAARNVLTSGRVAKIG 793  
QY 832 DFGLARDILSDSSYVVRGNARLPVKWMAPESLFEGIVTIKSDVMSYGIILLWEISLGNP 891



```

Db 794 DFGLARIMDSNVIKGNARLPVKWMAPEISFDCVTVQSDVMSYGILLWEISLGLNP 853
Qy 892 YPGIPVDANFYKLQSGFKMEQPYATEGIYFVMSQWAFDSRKRPSFNITSLFGCQLA 951
Db 854 YPGILVNSKFKYLVKDGQYMAQAPAFAPKNYISIMQACWALEPTRRPTFOQICSL--QKQ 911
Qy 952 EAEEACIRTSIHLPKQAA 969
Db 912 AQDRRVPNTNLPSSSS 929

RESULT 6
A49814
protein-tyrosine kinase (EC 2.7.1.112), receptor type kit precursor - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C/Accession: A49814; S49088
R/Tsujimura, T.; Hirota, S.; Nomura, S.; Nomura, M.; Tono, T.; Morii, E.; Kikuchi, T.
Blood 78, 1942-1946, 1991
A/Title: Characterization of Ws mutant allele of rats: a 12-base deletion in tyrosine kinase
A/Reference number: A49814; MUID:92003944; PMID:1912577
A/Accession: A49814
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-978 <TSU>
A/Cross-references: GB:D12524; NID:G220707; PIDN:BAA02094.1; PID:G220708
R/Tsujimura, T.; Tono, T.; Yamazaki, M.; Nomura, S.; Kitamura, Y.
submitted to the EMBL Data Library, October 1991
A/Description: Two isoforms of rat c-kit receptor tyrosine kinase.
A/Reference number: S49088
A/Accession: S49088
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-511, 516-978 <TS2>
A/Cross-references: EMBL:X62491; NID:G509135; PIDN:CAA44354.1; PID:G509136
C/Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;
C/Keywords: ATP; autophosphorylation; phosphotransferase; tyrosine-specific protein kinase
F/129-188/Domain: immunoglobulin homology <IM>
F/589-932/Domain: protein kinase homology <KIN>
F/597-605/Region: protein kinase ATP-binding motif

Query Match 23.4%; Score 1232.5; DB 1; Length 978;
Best Local Similarity 32.2%; Pred. No. 6.3e-56;
Matches 323; Conservative 176; Mismatches 340; Indels 165; Gaps 37;

Qy 36 CVLISHENSGSAGKPSYRMVRGSPDLQCTPRQSEGTVEAATVEVAESGSITLQVQ 95
Db 12 CVLILVLRGQTGTSQPSA-----SPEGPSPIQPAQSELIVEAGD-TIRLT 57

Qy 96 LATPGDLSCLVFKHSLGCGPHFDLQNRGIVSMALNV-TETQAGBYLLHIQSERANYT 154
Db 58 CTDPAPVK--WTFP-----ILLDVRIENKQSEWIR--EKAEATHT 92

Qy 155 VLFT-----VNVD--TOLYVLRPFRMENQDALLCISGEVPEPTV-EWVLC 200
Db 93 GKTYCVSGSLRGSIIYFVRDPAVLFLVLGLFLGKEDNDALVRC---PLTDPQVSNYSLI 149

Qy 201 SSHRESC-----KEGPAVRKEKVLHELPGTDIRCCARNALGRECTKLTIDNLQ 252
Db 150 ECDGKSLPTDLKFVNPFRAGITIKNRAVRLC---IRCAAQREGKWRMSDKFTLVKRA 206

Qy 253 A-----POSTLPQL--FLKVGPELMIRCKAIHNVHFGGLTW-----ELEDKALEE 295
Db 207 AIKAIPVSVPTSHLLKEGDTFTVICTIKDVSIVSDSMWIKLNPNQPSQAKVKNESWHQ 266

Qy 296 GSTFEMSTYTNRTMIRILIAFVSSVGRNDGTGYTCSKXHPQSALVT---ILEKGFIN 352
Db 267 GDF-----NYERQETLT-----ISSARVNDSGVPMCVANNTFGSANVTTLKVKVEGFIN 316

Qy 353 --ATSSQEEVEIDPYEKFCEFSVRKAYR-IRCTWIFSQASFPCEQGLE----DGYVIS 405
Db 317 IFPVKNTVFTVDG-ENVDLVVEFEAYPKPEHQOQWIYNNRT--PTNRGEDYVKSDNQSN 373

```

```

Qy 406 KFCDH-----KNKPGEYIFYAENDDAQPTMTFLNIRKYPQVLA--NASASQASCSDD 456
Db 374 RYVNELRLRLKGTGEGTYTFLVNSDVSAVTFDYYVNTKPEILTYDRLMNGRLQCQVAA 433
Qy 457 GYPLSPSWTKKCDKSPNCTEIEPE-GVWNKAKRKVFGOWVSSSTLANSEAGKLLVKC 515
Db 434 GPPEPTIDWYFCTGABQRCITVPVPDVOIQNASVSPFGKLVVQSSIDSSVFRHNGTVEC 493
Qy 516 CAYNSMGTSCETIFLNSPGPPFF-----IQDN-----ISFYATIGLCPLPFIWLIV 561
Db 494 KASNAVGKS--SAFFN---PAFKGNSKEQIQPHLFTPLLIQFVVTAGL---MGIIVM 543
Qy 562 LICHYKQKQRYESQLQMI-OVTGPLDNEYFVDFRDYEDYDLKWEFPRENLEFGKVLGSG 620
Db 544 VLAYKYLQKPMYEVQWKVEEING---NNYVYIDPTQLPYDHKWEFPRENLSFGKTLGAG 600
Qy 621 AFRGVNATAYGISKGTGVSIOVAVKMLKEKADSCKEKALMSLKMTHLGHHDHNVNLLG 680
Db 601 AFGKVEATAYGLIKSDAAMTAVAVMLKPSAHLTEREALMSLKVLSLGNHNMVNLIG 660
Qy 681 ACTLSGVPYLIIFYCCYCGDLNLYRSKREKFRHTWTETFEKHNFSSTSYPTFOAHSSNMPG 740
Db 661 ACTVGGPTLVITEYCCYCGDLNLYRKRDSF--IFSQEEQADAAALYKNLLHSKSSCDS 718
Qy 741 SREVQLHPPLDQLSGFN-----GNSIHSEDEIEYENOKRLAEDEEDLNVLTFEDL 791
Db 719 SNEY-----MDMKPGVSVVPTKTKRSARIDSYERDVTPTAIMEDDELALDL---EDL 770
Qy 792 LCPAYQVAKGMEFLEPKSVCHRDLAARNVLTHGVKVKICDFGLARDILSDSSVYVRGNA 851
Db 771 LSFYQVAKGWAPLASKNCIHRDLAARNILLTHGRITKICDFGLARDIRNDSNYVVKGNA 830
Qy 852 RLPVKWMAPESLPFGIYTIKSDVMSYGILLWEISLGVNVPYGPVDANFYKLQSGFKM 911
Db 831 RLPVKWMAPEISFNCVTFISDWSYSGIFLWELFSLGSSPYGPMVDSKFTKMLKEGPRM 890
Qy 912 EQPYATEGIYFVMSQWAFDSRKRPSFNITSLFGCQLABAE 955
Db 891 LSPHPAPAMYEVMVKTCDADPLKRPFTKQVQVLIQKISDSSK 934

RESULT 7
TVMSKT
protein-tyrosine kinase (EC 2.7.1.112), receptor type kit precursor - mouse
N/Alternate names: tyrosine kinase receptor c-kit
C/Species: Mus musculus (house mouse)
C/Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 31-Mar-2000
C/Accession: S00474; B44876; I49596
R/Qiu, P.; Ray, P.; Brown, K.; Barker, P.E.; Jhanwar, S.; Ruddle, P.H.; Besmer, P.
EMBO J. 7, 1003-1011, 1988
A/Title: Primary structure of c-kit: relationship with the CSF-1/PDGF receptor kinase
A/Reference number: S00474, MUID:88296403; PMID:2456920
A/Accession: S00474
A/Molecule type: mRNA
A/Residues: 1-975 <QIU>
A/Cross-references: GB:X00864; NID:G50423; PIDN:CAA68772.1; PID:G50424
R/Rossi, P.; Marziani, G.; Albanesi, C.; Charlesworth, A.; Gremia, R.; Sorrentino, V.
Dev. Biol. 152, 203-207, 1992
A/Title: A novel c-kit transcript, potentially encoding a truncated receptor, originate
A/Reference number: A44876; MUID:92331813; PMID:1378413
A/Accession: B44876
A/Molecule type: DNA
A/Residues: 771-814 <ROS>
A/Note: sequence extracted from NCBI backbone (NCBIN:108837, NCBI:108840)
R/Yasuda, H.; Galli, S.J.; Geisler, E.N.
Biochem. Biophys. Res. Commun. 191, 893-901, 1993
A/Title: Cloning and functional analysis of the mouse c-kit promoter.
A/Reference number: I49596; MUID:93221533; PMID:7682073
A/Accession: I49596
A/Status: preliminary; translated from GB/EMBL/DBD
A/Molecule type: DNA
A/Residues: 1-22 <RES>

```

A;Cross-references: GB:L11358; NID:g293325; PIDN:AAA37420.1; PID:g293326  
 C;Genetics:  
 A;Gene: kit; c-kit  
 A;Map position: 5  
 C;Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;  
 C;Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming protein;  
 fic protein kinase  
 F;1-22/Domain: signal sequence #status predicted <SIG>  
 F;23-975/Product: protein-tyrosine kinase kit #status predicted <KTC>  
 F;23-519/Domain: extracellular #status predicted <EXT>  
 F;51-100/Domain: immunoglobulin homology <IMM1>  
 F;130-189/Domain: immunoglobulin homology <IMM2>  
 F;220-295/Domain: immunoglobulin homology <IMM3>  
 F;331-397/Domain: immunoglobulin homology <IMM4>  
 F;426-496/Domain: immunoglobulin homology <IMM5>  
 F;520-542/Domain: transmembrane #status predicted <TM>  
 F;543-975/Domain: intracellular #status predicted <INT>  
 F;586-929/Domain: protein kinase homology <KIN>  
 F;594-602/Region: protein kinase ATP-binding motif  
 F;58-98.137-187,234-293,431-494/Disulfide bonds: #status predicted  
 F;146,296,303,323,355,370,466,489/Binding site: carbohydrate (Asn) (covalent) #status pr  
 F;622,639,790/Active site: Lys, Glu, Asp #status predicted  
 F;795,808/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 23.2%; Score 1223; DB 1; Length 975;  
 Best Local Similarity 32.1%; Pred. No. 1.9e-55;  
 Matches 324; Conservative 172; Mismatches 331; Indels 182; Gaps 39;

Qy 36 CVLISHENNGSAGKPPSYRMVGRGSPEDLOCTPRQSECTVVEATVEAESGSLTQVQ 95  
 Db 12 CVLLVLLRGQTATSPQSA-----SPGEPSPSHPAOSELIVEAGD-TL--- 54  
 Qy 96 LATPGDLISCL-----WVFKHSSILGQCPHF-----DLQNGRIVSMAILNVTTQAGEYLL 144  
 Db 55 -----SLTICIDPDFVRWTFK-----TYFEMVENKNEWIQEKA-----EATRTGTTC 98  
 Qy 145 HIQERANYTVLTVNVRD-TQLYVLRPPYPRKMNQDALL-----CIS 187  
 Db 99 ---SNSNGLTSSIVFVRDPKAKLFLVGLPLFGK-EDSDALVRCPLTPQVSNYSLEICDG 154  
 Qy 188 EG-----VPEPTVWVLCSSHRESCKEKGPAVVRKEKVLHFGTDIRCCARNALGR 240  
 Db 155 KSLPTDLTFVFPN-----KAGITIKNVKRAYH-----RLCVCAQAR 191  
 Qy 241 ECTKL-----FTIDLNOA-----POSTLPOL--FLKVGEPILWIRCKAIHNGHGFGLTW--- 286  
 Db 192 DGTWLHSDKFLTKREAKAPVSVVPETSHLLKGGDTFTVVCTIKDVSTSVNSWMLKNN 251  
 Qy 287 -----ELEDKALEEGSYFEMSTYSTNRITMIRILLAFVSSVGRNDGTGYTCSSSKHPSQ 339  
 Db 252 POPQHIAQVKNHSHRGDF-----NYERQETLT-----ISSARVDDSGVFMCYANNFTGS 301  
 Qy 340 SALVT---ILKGFNATS-SQEYEIDPYKFCFSVRKAYPR-IRCTWIPSOASPFCE 394  
 Db 302 ANVTTLTKVVEKGFNIPVKNVTTFVTDGENVDLVVEYAYPKPEHQOQWIMRT--SA 359  
 Qy 395 QRGLE-----DGVYSISKFDH-----KNKGEYIFAYENDDAQTKMFTLNIRKKPQVL 443  
 Db 360 NKGKDYVSKDNKSNIRYVQNLRLTKGTEGTYTFLVNSDASAVTFNVVNTKPBIL 419  
 Qy 444 A--NASASQASCSGYPPLSPWTKKCSKSPNCTEEI-PEGVMNKNRKFVFGQWVSS 500  
 Db 420 TYDRLINGMLQVAAEGPPEPTIDWYFCTGAEQRCTTPSPVDVQVQVSVSPFGKLVQS 479  
 Qy 501 TLNMBEAGKGLLVKCCAYNSMGTSCTETPLNSPGPPFIQNISFYATIGLCPLPIV--- 557  
 Db 480 SIDSSVFRHNGTVECKASNDVGKS--SAFFN-----FAFKEQIQANTLFTPLLLIGFWAAG 533  
 Qy 558 ---VLIVLICHYKQKPPYEQSLOMI-QVTGPDNVEYFVDRDYEDLKWEEPRENLEF 613  
 Db 534 AMGIVLWLTYYTKYLOKPMYEQVQWKEEING-----NNYVIIDTQLPYDHRKWEFFRNLSE 590  
 Qy 614 .GKVLGSGAFGRVMNATAYGISKTVGSIQVAVKMLKEKADSCKEALMSLKMWTLGHHD 673

RESULT 8

TVHUMD

N;Contains: protein-tyrosine kinase (EC 2.7.1.112) csflr/fms  
 C;Species: Homo sapiens (man)  
 C;Date: 28-Dec-1987 #sequence revision 31-Dec-1993 #text change 16-Jul-1999  
 C;Accession: S08123; A24533; I56672; I57648; I59083; I52772  
 R;Hampe, A.; Shamoon, B.M.; Gobet, M.; Sherr, C.J.; Galibert, F.  
 Oncogene Res. 4, 9-17, 1989  
 A;Title: Nucleotide sequence and structural organization of the human FMS proto-oncogen  
 A;Reference number: S08123; MUID: S9239490; PMID: 2524025  
 A;Accession: S08123  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-972 <HAM>  
 A;Cross-references: GB:U63963; EMBL:X14720; NID:g1915975; PIDN:AAB51696.1; PID:g1915976  
 A;Note: this sequence was submitted to the EMBL Data Library, March 1989  
 R;Cousens, L.; Van Beveren, C.; Smith, D.; Chen, E.; Mitchell, R.L.; Isacke, C.M.; Ver  
 Nature 320, 277-280, 1986  
 A;Title: Structural alteration of viral homologue of receptor proto-oncogene fms at car  
 A;Reference number: A24533; MUID: 86175013; PMID: 2421165  
 A;Accession: A24533  
 A;Molecule type: mRNA  
 A;Residues: 1-53, 'A', 55-972 <COU>  
 A;Cross-references: GB:J03149  
 A;Note: the authors translated the codon GCA for residue 54 as Pro  
 R;Wheeler, E.F.; Rousael, M.F.; Hampe, A.; Walker, M.H.; Fried, V.A.; Look, A.T.; Rette  
 J. Virol. 59, 224-233, 1986  
 A;Title: The amino-terminal domain of the v-fms oncogene product includes a functional  
 sequences.  
 A;Reference number: I56672; MUID: 86281820; PMID: 3525854  
 A;Accession: I56672  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-16 <RES>  
 A;Cross-references: GB:M14002; NID:g182676; PIDN:AAA35849.1; PID:g553292  
 R;Visvader, J.; Verma, I.M.  
 Mol. Cell. Biol. 9, 1336-1341, 1989  
 A;Title: Differential transcription of exon 1 of the human c-fms gene in placental troph  
 A;Reference number: I57648; MUID: 89261741; PMID: 2524648  
 A;Accession: I57648  
 A;Status: preliminary  
 A;Molecule type: translated from GB/EMBL/DBJ  
 A;Residues: 1-16 <RE2>  
 A;Cross-references: GB:M25706; NID:g349454; PIDN:AAA58421.1; PID:g553224  
 R;Browning, P.J.; Bunn, H.F.; Cline, A.; Shuman, M.; Nienhuis, A.W.  
 Proc. Natl. Acad. Sci. U.S.A. 83, 7800-7804, 1986  
 A;Title: Replacement of COOH-terminal truncation of v-fms with c-fms sequences marked  
 A;Reference number: I59083; MUID: 87017034; PMID: 3532121

A;Accession: 159083  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 874-972 <RE3>  
A;Cross-references: GB:M14193; NID:g182521; PIDN:AAA35834.1; PID:g182522  
R: Nienhuis, A.W.; Bunn, H.F.; Turner, P.H.; Gopal, T.V.; Nash, W.G.; O'Brien, S.  
Cell 42, 421-428, 1985  
A;Title: Expression of the human c-fms proto-oncogene in hematopoietic cells and its del  
A;Reference number: 152772; MUID:85282599; PMID:4028159  
A;Accession: 152772  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 244-295 <RE4>  
A;Cross-references: GB:M11067; NID:g182674; PIDN:AAA35848.1; PID:g442423  
C;Genetics:  
A;Gene: GDB:CSF1R; FMS  
A;Cross-references: GDB:120600; OMIM:164770  
A;Map position: 5q33.2-5q33.3  
A;Introns: 17/1; 103/1; 198/1; 243/3; 297/1; 361/2; 400/1; 440/2; 504/1; 542/3; 585/1; 6  
C;Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;  
C;Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming protein;  
fic protein kinase  
F;1-23/Domain: signal sequence #status predicted <SIG>  
F;24-972/Product: macrophage colony-stimulating factor 1 receptor #status predicted <MAT  
F;24-512/Domain: extracellular #status predicted <EXT>  
F;35-86/Domain: immunoglobulin homology <IMW1>  
F;120-179/Domain: immunoglobulin homology <IMW2>  
F;217-280/Domain: immunoglobulin homology <IMW3>  
F;316-383/Domain: immunoglobulin homology <IMW4>  
F;412-487/Domain: immunoglobulin homology <IMW5>  
F;513-537/Domain: transmembrane #status predicted <TMW>  
F;538-972/Domain: intracellular #status predicted <INT>  
F;580-917/Domain: intracellular #status predicted <INT>  
F;588-596/Region: protein kinase ATP-binding motif  
F;42-84,127-177,224-278,419-485/Disulfide bonds: #status predicted  
F;45,73,153,240,275,302,335,353,412,428,480/Binding site: carbohydrate (Asn) (covalent)  
F;616,633,778/Active site: Lys, Glu, Asp #status predicted  
F;783,796/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 23.2%; Score 1221.5; DB 1; Length 972;  
Best Local Similarity 32.5%; Pred. No. 2.3e-55;  
Matches 330; Conservative 152; Mismatches 331; Indels 201; Gaps 33;

QY 76 VYEAATVEVAESGSIITLOVQLATPGDLSCLWFKHSLGCGPHDQNRGIVSMAIL--- 132  
DB 22 VIEPSPPELVVKGAT--VTLRGVNGSGVENDGPPS-----PHWTLYSDG--SSSILSTN 72

QY 133 NVTETQAGEYLLHTQSERANYTVLFTVNVDRDTQLYVLRPY-----FRKMNQDALL-C 185  
DB 73 NATFQNTGTVRCTEFGDPLGSSRAIHLYVKDP-----ARPWNVLAEVVFEDQDALLPC 127

QY 186 ISGVPEPTVEVWVLCSSHRESCKEGBPAVVRKEEK-----VLHE---LF 226  
DB 128 L--LTDPLV-----EAGVSLVVRGRPLMRHTNYSFSPWHGFTIHRAKFIQ 171

QY 227 GTDIRCCARNALGRECTKLTIDLNQAPQSTLP-----QLFKVGEPLWIRKCAI 276  
DB 172 SDQYQCSALMG-GR---KWSISIRLKVQKVIQPGPALTILVPAELVRAIRGAAQIVCSAS 227

QY 277 HVNHGFGLTWELEDKALEEGSYFEMSTYSNTRMIRILLAFVSVSGRNDGYITCSS-- 334  
DB 228 SVDVNFDFVLOHNNTKL---AIPQSDFNHNRVQ-KVLTLNLDQVDPQHAGNYSVCASNV 283

QY 335 --KHPQSQALVTILEKGFINATSSQE-EYBIDPYEKFCSVRPKAYPRIR-CWIFSQAS 390  
DB 284 QGKH-STSMFPRVVEAYLNLSEQNLIQEVTVGEGNLKVMVEAYPGLGFNWTY----- 338

QY 391 FPCQRLGLEDGYSIKFCDHNKP-----GYIFYAEN 423  
DB 339 -----LGPFSHQPEPKLANATTKDTRHTFTLSLPLKPSAGRYSFLARN 385

QY 424 DDAQFTKMTLNRKKPOV-----LANASASQACSSDGYPLPSWTWKKCSKSPNCTEE 478

DB 386 PGMRALTFELTLRYPPEVSVIWTFINSGT-LLCAASGYPPQPNVTWLCOSGHTDRCDRA 444  
QY 479 IPEGVMN-----KCANRKVFQGMVSSSTLNMSEAGKGLLVKCCAYNSMGTSCTETFLNSFG 534  
DB 445 QVLQVMDPPYEVLSQSEPFHKVTVQSLTTLVETLEHNTVECKRAHNSVSGSWAFIPISAG 504  
QY 535 PPFQIDNISFYATIGLC--LPFIVLVILVICHKKQFQRYESQLOMIQVTPGLDNEYF 591  
DB 505 ATHPPDEFFLTPVVVACMSIMALLLLLLLYKKQPKYQVVRWKIIE--SYEGNSY 562  
QY 592 YVDFRDYEDLKWEFPFRENLEFGVLGSGAFGRVVMNATAGYISKTGVSIOVAVKMLKEKA 651  
DB 563 PIDTQPLPYNEKWEFPFNNLOFGKTLGAGAFQKVVEATAFGLGKEDAVLKVAVRMLKSTA 622  
QY 652 DSCKEALMSELKQWTHLGHHDNINVLGACTLGGPVYLI FEYCCYGGDLLNLYRSKREKF 711  
DB 623 HADEKEALMSELKIMSHLQGHENIVNLLGACTHGGPVLVITEYCCYGGDLLNLYRSKREKF 680  
QY 712 HRTWTEIFKEHNFSSYTFQAHSSNMPGSRREVQLHPDLQSLGPNNGNSIHSEDE----- 766  
DB 681 -----AMLGP---SLSPQDPEGGVDYKNIHLEKYYVRD 712

QY 767 -----TEYENQKELAE--EEDLNVLTFEDLLCFAYOVAKGMEFLBPK 808  
DB 713 SGFSSQGVDTYVEMRPVSTSNDSFSQDDDKDGRPLELRDLHLHFSSQVAGMAFLASK 772

QY 809 SCVHRDLAARNVLVTHGKVKIKCDFGLARDILSDSSYVVRGNARLPVKWMAPESEFEGY 868  
DB 773 NCIRHDAARNVLTNGHAKIGDFGLARDIMDSNIVKGNARLPVKWMAPESEFIDCVY 832

QY 869 TIKSDVMSYGILLWEIFSLGVNYPGIPVDANFYKLIQSGFPMEQPFYATEGIVFMQSC 928  
DB 833 TVQSDVMSYGILLWEIFSLGVNYPGIPVDANFYKLIQSGFPMEQPFYATEGIVFMQSC 892

QY 929 WAFDSRKRPFPNLTSLGQLAEAEACTRTSHLPKQAAPOQRGRLAQSPO 982  
DB 893 WALEPHTHPTFOQICSEFL--QEQAQEDRRERDYNLFPSS---RSGGSGSSSE 941

RESULT 9  
151703  
C-kit-related kinase 1 (XKrk1) - African clawed frog  
C;Species: Xenopus laevis (African clawed frog)  
C;Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 18-Jun-1999  
C;Accession: I51703  
R;Baker, C.V.; Sharpe, C.R.; Torpey, N.P.; Heasman, J.; Wylie, C.C.  
Mech. Dev. 50, 217-228, 1995  
A;Title: A Xenopus c-kit-related receptor tyrosine kinase expressed in migrating stem c  
A;Reference number: I51703; MUID:95344996; PMID:7619732  
A;Accession: I51703  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-954 <BAK>  
A;Cross-references: EMBL:Z48770; NID:g763033; PIDN:CAA8688.1; PID:g763034  
C;Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology  
C;Keywords: ATP  
F;575-915/Domain: protein kinase homology <KIN>  
F;583-591/Region: protein kinase ATP-binding motif

Query Match 23.2%; Score 1220; DB 2; Length 954;  
Best Local Similarity 32.0%; Pred. No. 2.7e-55;  
Matches 315; Conservative 160; Mismatches 321; Indels 188; Gaps 35;

QY 77 YEAAATVEVAESGSIITLOVQLATPGDLSCLWFKHSLGCGPHDQNRGIVSM 130  
DB 15 YTGDAVPKINDGEDRVTVNVGDKVSLCEDAHLVTLAFQKSLMKKPR-DLKSRLP----- 68

QY 131 ILNVTET-----QAGEYLLHIQSERANYTVLFTVNVDRDTQLYVLRPYPRKME 178  
DB 69 -LNNSETDOFFVILKADLRHIGYICTNTETQNTSV--SLFVKDPAFPDIPFDVTE 125

QY 179 NQDAL-LCISEGVPEPTVEVWVLCSSHRESCKEGBPAVVRKEEKVLHFLFTDIRC----- 232



RESULT 11  
S16385  
macrophage colony-stimulating factor 1 receptor precursor - rat  
N:Contains: protein-tyrosine kinase (EC 2.7.1.112) CSF-1R  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 02-Jun-2000  
C:Accession: I60321; S16385  
R:Borycki, A.G.; Guillier, M.; Leibovitch, M.P.; Leibovitch, S.A.  
Growth Factors 6, 209-219, 1992  
A:Title: Molecular cloning of CSF-1 receptor from rat myoblasts. Sequence analysis and  
A:Reference number: I60321; MUID:93001225; PMID:1389227  
A:Accession: I60321  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-978 <RES>  
A:Cross-references: EMBL:X61479; NID:957543; PID:CAA43706.1; PID:G57544  
A>Note: in Genbank entry RRGSF1, release 113.0, the source is designated as Rattus rattus  
A:Note: submitted to the EMBL Data Library, August 1991  
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;  
C:Keywords: ATP; autophosphorylation; glycoprotein; growth factor receptor; kinase-related  
protein; tyrosine-specific protein kinase  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-978/Product: macrophage colony-stimulating factor 1 receptor #status predicted <MAI>  
F:20-515/Domain: extracellular #status predicted <EXT>  
F:35-86/Domain: immunoglobulin homology <IMM1>  
F:120-179/Domain: immunoglobulin homology <IMM2>  
F:217-280/Domain: immunoglobulin homology <IMM3>  
F:316-381/Domain: immunoglobulin homology <IMM4>  
F:410-485/Domain: immunoglobulin homology <IMM5>  
F:516-535/Domain: transmembrane #status predicted <TM>  
F:536-978/Domain: intracellular #status predicted <INT>  
F:578-915/Domain: protein kinase homology <KIN>  
F:586-594/Region: protein kinase ATP-binding motif  
F:42-84,127-177,224-278,417-483/Disulfide bonds: #status predicted  
F:45,73,302,335,389,410,449,478,491/Binding site: carbohydrate (Asn) (covalent) #status  
F:614,631,776/Active site: Lys, Glu, Asp #status predicted  
F:781,794/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 23.0%; Score 1210; DB 2; Length 978;  
Best Local Similarity 33.5%; Pred. No. 8,9e-55;  
Matches 318; Conservative 147; Mismatches 329; Indels 154; Gaps 31;

QY 78 EAATVEVAESGIIITQVQLATPDLSCLVFKHSIGCQHPDLQNRGVSMALNVTTET 137  
DB 36 EFTVLRCSVNGSVEWD-----GPSFYWTLDPSGSGS-----TLTRNATFK 77  
QY 138 QAGEY-LLHIQGERANYTVLFTVNVEDTOLYV-----HLVYKDPAHSNWLLAQEVTVVGEQEAFLPCL- 128  
DB 78 NTGYTRCTELEDPMAGSTTI-----LRRPYFRKMNQDALL-CIS 187  
QY 188 EGVPEPTVEWVLCSSHRESCKEKGPAVVRK-----BEKVLHFLGTDIRCC 233  
DB 129 --ITDPALK-----DSVSLMREGGRQLKTYFFSAWGRGFIRKAKVL-----DSNTYVC 177  
QY 234 ARNALGRECTKL-FTIDLN-----QAPQSTL-POLFLKV-GEPLWIRCKAIHNVHGFGFTW 286  
DB 178 KTWNGRESTSTGIWLKVRNVHPEPPQIKLEPSKLVIRGEAAQIVCSATNAEVEGFNVIL 237  
QY 287 ELEDKALE--EGSYEMSTYSTNRTWIRILLAFVSVSGVNDGYTC-SSSKHPQSALV 343  
DB 238 KRQDTKLEIPLNSDFQDNYKKVRAL-----SLNAVDPQDAGIYCSVASNDVGRTATM 291  
QY 344 T--ILEKFINATSSQE-EYEIDPYEKFCSFRKAYRPRIR-CTWIFSQASPPCORGLE 399  
DB 292 NFQVVESAYVNLTSQSLQLQSVSGDSLILTVHADAYPSIQHNTYTLGPFFE-DORKLE 350  
QY 400 -----DGYISIKPCDH--KNKPGEIFYAEINDDAQFTKMTLIRKKPQVLANASASQ 450  
DB 351 FITORAIYRYTFKFLNVRKASEAQYFLMAQNKAGNNLTFLTLRYPPPEVSVTWMPVN 410  
QY 451 AS-----CSSDGYPLPSWTWKKCKSKDPNCTBEIPGVWN-----KKANRKRFGQWVSSSTL 502

DB 411 GSDVLFCDVSGYPOPSVTWMBECRGHTDRCDQAALQVWMDTPEVLSQKPFDFKVIQSQL 470  
QY 503 NMSEAGKGLLVKCAYNAMSGTSCETIFLNSPGFFPFQIDONISFYATIGLC---LPIVLVL 559  
DB 471 PIGTLKHNMTYFCKTHNSVGNSSQYFRAVLSQSQQLPDESFTPPVVAACMSVMSLVL 530  
QY 560 IVLICHKYYKQFVYESQLQMIQ-VTGPLDNEYFYVDPRDYEDLKWEPFRENLEFGKVLG 618  
DB 531 LLLLLLYKQKPKQVRWKIIERYEG---NSYTFIDPTQLPYNKEWEPFNNLQFQKTLG 587  
QY 619 SGAGRWNNATAYGISKTCVSIQAVAKLKEKADSCKEALMSELKMMTHLGHHDNIVNL 678  
DB 588 AGAFGKVVETATFGLGKREDAVLKAVAKLKEKADSCKEALMSELKMMTHLGHHDNIVNL 647  
QY 679 LGACTLSGPVYLIPEYCCYGLLNLVRSKRE-----KFRHTWTBI 718  
DB 648 LGACTHGGPVLVITYEYCYGLLNLVRSKRE-----KFRHTWTBI 707  
QY 719 FKEHNFSSYPTFOAHSSNMSPGSRREVQLHPPLDQLSGFNGSIHSEDEIEYENQKRLAE 778  
DB 708 RRDGSGFS-----SQGVDTYVEMRPVSTSSSDSFFKQD-----L 740  
QY 779 EEDLNVLTEDLLCFAYQVAKGMEFLPKSCVHRDLAARNVLVTHGKVKVIXCDFGLARD 838  
DB 741 DKPEFRPLELDLWDLHFSSQVACQNAFLASKNCIHRDVAARNVLLTSGHVAKIGDFGLARD 800  
QY 839 ILSDSYVVRGNARLPVKWMAPELSLFGIYTIKSDVMSYGLLWEIFSLGVNPNYPGIPVD 898  
DB 801 IMDSNVVVRGNARLPVKWMAPELSLFGIYTIKSDVMSYGLLWEIFSLGVNPNYPGILVN 860  
QY 899 ANFYKLTQSGFKMEQPPYATYEGYFVVMQSWAFDPRKRPSFNLTSFL 946  
DB 861 NKYKLVKDGQVMAQPVFAPAKNIYSINQSCWDLPEPTRPTQQICFL 908

RESULT 12  
TVMSMD  
macrophage colony-stimulating factor 1 receptor precursor - mouse  
N:Contains: protein-tyrosine kinase (EC 2.7.1.112) csf1r/fms  
C:Species: Mus musculus (house mouse)  
C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 02-Jun-2000  
C:Accession: S01880  
R:Rothwell, V.M.; Rohrschneider, L.R.  
Oncogene Res. 1, 311-324, 1987  
A:Title: Murine c-fms cDNA: cloning, sequence analysis and retroviral expression.  
A:Reference number: S01880; MUID:88217329; PMID:2966922  
A:Molecule type: mRNA  
A:Residues: 1-976 <ROT>  
A:Cross-references: EMBL:X06368  
C:Genetics:  
A:Gene: fms  
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;  
C:Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming protein;  
f1c protein kinase  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-976/Product: macrophage colony-stimulating factor 1 receptor #status predicted <MAI>  
F:20-515/Domain: extracellular #status predicted <EXT>  
F:35-86/Domain: immunoglobulin homology <IMM1>  
F:120-179/Domain: immunoglobulin homology <IMM2>  
F:217-280/Domain: immunoglobulin homology <IMM3>  
F:316-381/Domain: immunoglobulin homology <IMM4>  
F:410-485/Domain: immunoglobulin homology <IMM5>  
F:516-535/Domain: transmembrane #status predicted <TM>  
F:536-976/Domain: intracellular #status predicted <INT>  
F:578-914/Domain: protein kinase homology <KIN>  
F:586-594/Region: protein kinase ATP-binding motif  
F:42-84,127-177,224-278,417-483/Disulfide bonds: #status predicted  
F:45,73,302,335,389,410,449,478,491/Binding site: carbohydrate (Asn) (covalent) #status  
F:614,631,776/Active site: Lys, Glu, Asp #status predicted  
F:781,794/Binding site: magnesium (Asn, Asp) #status predicted







Db 353 -----TDHRWHTPTSPSTMQEHI YHARLQKRMNAQE 385  
Qy 414 PGEYFYAENDDAQTKMTLNIRKKP-QVLNANASQASCSGDGVPPLSPMTWKKCSDKS 472  
Db 386 QOQYTFYAKSNLANGSI SFHVKNYQKPTAVVWENHITLTCTSGYPAPQIWIYQCSGR 445  
Qy 473 PNCT-----BEIPEGVWNNKANRVPQGVQVSSSTLMSBAGKGLLVKCCAYNSMGTS 524  
Db 446 FTCNGNNTGLPKQNHPOAL-TVEVQREYGAVEVESFTVGLSNHRMTVECVAFNLVGV 504  
Qy 525 CETIFLSPGPPPTQDNISFATIGLC-LPFIIVLIVLCHYKQKQRYESQLOMIOV 582  
Db 505 SDTFTVE-----VSDKLTFTLGAAGVLAIFLLLVFLYKQKQKQPFIRKWIIEA 557  
Qy 593 TGPLDNEFYVDFRDEYDLKWEFPRENLEPGKVLGSGAFGRVNNATAYGISK-TGVSIQ 641  
Db 558 RE--GNNTFTDPTQLPYNEKEWFPKDLKGLKVLGAGAFGVKVEATAPGLGEDKXNLTIR 615  
Qy 642 VAVKMLKEKADSCKEALMSELKMMTHLGHHDNI VNLGACTLSPGVVLIPIEYCCYGDLL 701  
Db 616 VAVKMLKANAHSDEREALMSELKILSHLGHQNI VNLGACTYGGVVLVITEYCSLGDLL 675  
Qy 702 NYLRSKREKPHRTWTEIPEKHNFSYPTQAHN-----SSMPGSR 743  
Db 676 NFLROKAETFFVNLVNNI-----PEIMENSNDYKNI CNQKYIRSDSGISSTSTSY 726  
Qy 744 VOLHPLQLSGFNGSIHSEDEIEVENOKRLAEBEEDLNLVTPEDLLCPAYQVAKGME 803  
Db 727 LEMRP-----SQSHIIEASGRKSLCEBNGD--WFLDIDDLRFLSLQVAGLD 771  
Qy 804 FLEPKSCVHRDLAARNVLVTHGKVKIKDPGLARDILSDSSVYVRGNARLPVKWMAPESL 863  
Db 772 FLASRNCIHRDVAARNVLLTDKRAKIKDPGLARDIMDSNVYVKGARLPVKWMAPESI 831  
Qy 864 PEGYITIKSDVWSYGILLWEIPSLGVPYGPVDPANFKLQSGFKMEOPPYATEGIFY 923  
Db 832 FDCVYTVQSDVWSYGILLWEIPSLGKSPYMAVDSPRYKMKRGYQMSQDPFALPEIYM 891  
Qy 924 VMQSCWAFDSRRKPPNLTSLFLGCOLA--EAEACIRTSIHLPKOAAPOQRGLRAQSP 981  
Db 892 IMKMCWNLEPTERTPTFSMISQMINRLGGQDEKLIYRNVQ-PEQVAE-----GEACDEP 946  
Qy 982 QR 983  
Db 947 KR 948

RESULT 15  
JN0677  
protein-tyrosine kinase (EC 2.7.1.112), receptor type kit precursor - chicken  
N:Alternate names: tyrosine kinase receptor kit  
C:Species: Gallus gallus (chicken)  
D:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Jun-2000  
C:Accession: JN0677  
R:Sasaki, E.; Okamura, H.; Chikamune, T.; Kanai, Y.; Watanabe, M.; Naito, M.; Sakurai, M.  
Gene 128, 257-261, 1993  
A:Title: Cloning and expression of the chicken c-kit proto-oncogene.  
A:Reference number: JN0677; PMID:93292995; PMID:7685729  
A:Accession: JN0677  
A:Molecule type: mRNA  
A:Residues: 1-960 <SAS>  
A:Cross-references: DBJ:D13225; NID:G303532; PIDN:BAA02506.1; PID:G303533  
A:Experimental source: brain  
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;  
C:Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming protein;  
protein kinase  
F:1-24/Domain: signal sequence #status predicted <SIG>  
F:25-960/Product: tyrosine kinase receptor #status predicted <MAT>  
F:314-380/Domain: immunoglobulin homology <IMM>  
F:573-916/Domain: protein kinase homology <KIN>  
F:581-589/Region: protein kinase ATP-binding motif  
F:76,135,149,269,286,306,318,338,356,453,469/Binding site: carbohydate (Asn) (covalent)

Query Match 22.4%; Score 1181; DB 1; Length 960;  
Best Local Similarity 31.8%; Pred. No. 2.7e-53;  
Matches 310; Conservative 159; Mismatches 356; Indels 150; Gaps 30;  
Qy 45 GSSACKPSRYMRVGRSPEDLOCTPRQSEGTVEAATVEAESGSITLQVQLATPDLSC 104  
Db 24 GSVPHSESSLVNVKGEELRKN-----EEGPTVMNFQNSDPSAKTR 65  
Qy 105 L-----WPKHSSLCQPHFQNLQNRGIVSMALNVTTETQAGEYLLHIQSERANYTVLFTV 159  
Db 66 ISNEKEWTKNATIRIDIGRYECKSKG-----SIVN-----SFIV 99  
Qy 160 NVRDQLVLRPPYPRKMNODALLCISEGVEPTEVWVLCSSHRESCKEKGAVVRKEE 219  
Db 100 FKVDPNVFLVDSLTYGKEDSDILLVCLTPD-VLNFTRKCDGKPLPKNMTFIPNPK 158  
Qy 220 ----KVLHELFGTDIRCCAR-NALGRECTKLTIDLNOAP-QSTLPOL-----FLKV 266  
Db 159 GIIHKVQSRFKGCYQCLAKHNGVEKISEHIP--LNVRFVHKALFVITLSKSYELLKEG 215  
Qy 267 EPLWIRKAIHVNHGFGLTWELEDKALEEGSYFEMSTYSTNRTMIRILLAFVSSVGRNDT 326  
Db 216 BEFEVTCIITDVS SVKASWISYKSAIVTSKRNIGDYGERK----LTLNRSVGND 271  
Qy 327 GYITCSSSKHP--SQSALVTI--LEKGFNATSSQE-EYEDIPYEFKFCFSVRKAYPRIR 381  
Db 272 GEPTC-QAENPFKTNATVTLKALAKGFVRLFATMTTIDINAGQGNLTVEYAYPK-- 328  
Qy 382 CTWIFSOASFPCEQRLGLEDGYSISKFDHKNK-----PGEYIF 419  
Db 329 -----PKEEYVMYMNELQNSDHYKFKTVGNNSYTSSELHLTRLKGTBEGIYF 378  
Qy 420 YAENDDAQTKMTLNIRKKQVLANASAQ--ASCSSDGVPLSPMTWKKCDKSPNCTE 477  
Db 379 FVNSDASSSTVFNVYKTKPEILLTDLMLNDILQCVATGPPAPTIYWFYFCGEORCLD 438  
Qy 478 E---IPEGVWNNKANRKV--FCQWSSSTLANSEAGKGLLVKCCAYNSMGTS CETIFLS 532  
Db 439 SPTISPMVKVSYTNSSVPSFERILVESTVNAS-MFKSTGTICCEASSNGDK-SSVFFNF 496  
Qy 533 PGPPFFIQDNISFYA-----TGLCLPFIIVLIVLCHYKQKQRYESQLOMI-QVTG 584  
Db 497 A-----IKEQIRTHLTFTPLLIAGVAAAGLMCIIVMILVYIYLOPKYEVQWKVVEEING 551  
Qy 585 PLDNEYFYVDFRDEYDLKWEFPRENLEPGKVLGSGAFGRVNNATAYGISKTVSIQAV 644  
Db 552 ---NNVYIIDPTQLPYDHKWEFPNRLSFGKTLGAGAFGVKVEATAPGLKSDAAMTVAV 608  
Qy 645 KMLKEKADSCKEALMSELKMMTHLGHHDNI VNLGACTLSPGVVLIPIEYCCYGDLLNYL 704  
Db 609 KMLKPSAHLTEREALMSELKVLVSLGNHINIVNLGACTIGGPTLVITEYCCYGDLLNLF 668  
Qy 705 RSKREKP-----HRTWTE--IFKEHNFSSYPTQA-HSNSSMPGSRVQLHPPLDQLSGFN 757  
Db 669 RKRKDSFICPKHEEHAEEAVYENLLHQAEPTADAVNEYMDMKPGVSYAVPPKADKKRPVK 728  
Qy 758 GNSIHSEDEIEVENOKRLAEBEEDLNLVTPEDLLCPAYQVAKGMEFLEPKSCVHRDLAA 817  
Db 729 SGS-----YTDQDVTLSMLEDDALDEDLDSFYSQVAKGMSFLASKNCHIRDLAA 780  
Qy 818 RNVLVTHGKVKIKDPGLARDILSDSSVYVRGNARLPVKWMAPESLFEGIYTIKSDVWSEY 877  
Db 781 RNILLTHGRITIKDPGLARDIRDSNVYVKGARLPVKWMAPESFNCVYTFESDWSY 840  
Qy 878 GILLWEIPSLGVPYGPVDPANFKLQSGFKMEOPPYATEGIFYVWQSWAFDSRRKRP 937  
Db 841 GILLWEIPSLGSSPYGMPVDSKPYKMIKEGRVMPSPSPPEMYDIMKSCWDADLPQRP 900  
Qy 938 SFPNLTSLFLGCOLAE 952  
Db 901 TFKQIVQLIEQQLSD 915

Search completed: August 26, 2003, 07:29:00  
Job time : 36.9834 secs

---



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2003, 07:14:08 ; Search time 18.9904 Seconds  
(without alignments)  
2456.527 Million cell updates/sec

Title: US-09-919-408A-2  
Perfect score: 5264  
Sequence: 1 MRALQSRDRRLLEWLSV.....RGLRAQSPQVKIHRS 992

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5264	100.0	992	1 FLT3 MOUSE	Q00342 mus musculus
2	4429.5	84.1	993	1 FLT3 HUMAN	P36888 homo sapien
3	1258.5	23.9	980	1 KFN5_FELCA	P13369 felis silve
4	1249.5	23.7	978	1 KFN5_RSMD	P00545 feline sarc
5	1223.5	23.2	975	1 KIT MOUSE	P05532 mus musculus
6	1222.5	23.2	978	1 KIT CAPHI	Q28317 capra hircu
7	1221.5	23.2	972	1 KFN5 HUMAN	P07333 homo sapien
8	1219	23.2	977	1 KFN5_MOUSE	P09581 mus musculus
9	1215	23.1	977	1 KIT BOVIN	P43481 bos taurus
10	1210	23.0	978	1 KFN5 RAT	Q00495 rattus norv
11	1207	22.9	975	1 KIT CANFA	O97799 canis famil
12	1196	22.7	976	1 KIT HUMAN	P10721 homo sapien
13	1181	22.4	960	1 KIT CHICK	Q08156 gallus gall
14	1178.5	22.4	978	1 KIT_FELCA	Q28889 felis silve
15	1157.5	22.0	1088	1 PGDS RAT	P20786 rattus norv
16	1152.5	21.9	1089	1 PGDS_HUMAN	P16234 homo sapien
17	1142.5	21.7	1089	1 PGDS_MOUSE	P26618 mus musculus
18	1132	21.5	1087	1 PGDS_XENLA	P26619 xenopus lae
19	1098	20.9	1098	1 PGDR_MOUSE	P05622 mus musculus
20	1078	20.5	1106	1 PGDR_HUMAN	P09619 homo sapien
21	998.5	19.0	1338	1 VGR1_HUMAN	P17948 h vascular
22	984.5	18.7	1336	1 VGR1_RAT	P53767 rattus norv
23	983.5	18.7	1333	1 VGR1_MOUSE	P35969 mus musculus
24	967.5	18.4	370	1 KIT_FSVHZ	P04048 feline sarc
25	965	18.3	1348	1 VGR2_COTJA	P52583 coturnix co
26	961.5	18.3	1356	1 VGR2_HUMAN	P35968 homo sapien
27	954.5	18.1	1298	1 VGR3_HUMAN	P35916 homo sapien
28	952	18.1	1363	1 VGR3_MOUSE	P35917 mus musculus
29	934	17.7	1343	1 VGR2 RAT	O08775 rattus norv
30	926.5	17.6	1367	1 VGR2_MOUSE	P35918 mus musculus
31	818	15.5	813	1 FGR2_XENLA	Q03364 xenopus lae
32	810.5	15.4	821	1 CEK3_CHICK	P18461 gallus gall
33	791	15.0	823	1 FGR2_HUMAN	P21802 homo sapien

ALIGNMENTS

RESULT 1				
FLT3_MOUSE				
ID	FLT3_MOUSE	STANDARD;	PRT;	992 AA.
AC	Q00342;			
DT	01-APR-1993 (Rel. 25, Created)			
DT	01-APR-1993 (Rel. 25, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	FL cytokine receptor precursor (EC 2.7.1.112) (Tyrosine-protein kinase receptor flk-2) (Fetal liver kinase 2) (Tyrosine-protein kinase FLT3).			
DE	receptor flk-2) (Fetal liver kinase 2) (Tyrosine-protein kinase FLT3).			
GN	FLT3 OR FLT-3 OR FLK-2.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
RX	MEDLINE=91292518; PubMed=1648448;			
RA	Matthews W., Jordan C.T., Wiegand G.W., Pardoll D., Lemischka I.R.;			
RT	"A receptor tyrosine kinase specific to hematopoietic stem and progenitor cell-enriched populations.";			
RL	Cell 65:1143-1152(1991).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92019834; PubMed=1656368;			
RA	Rosnet O., Marchetto S., Delapeyriere O., Birnbaum D.;			
RT	"Murine Flc3, a gene encoding a novel tyrosine kinase receptor of the PDGFR/CSFIR family.";			
RL	Oncogene 6:1641-1650(1991).			
RN	[3]			
RP	CHARACTERIZATION.			
RX	MEDLINE=93205405; PubMed=8384358;			
RA	Maroc N., Rottapel R., Rosnet O., Marchetto S., Lavezzi C., Mannoni P., Birnbaum D., Dubreuil P.;			
RT	"Biochemical characterization and analysis of the transforming potential of the FLT3/PLK2 receptor tyrosine kinase.";			
RL	Oncogene 8:909-918(1993).			
CC	-!- FUNCTION: RECEPTOR FOR THE FL CYTOKINE. HAS A TYROSINE-PROTEIN KINASE ACTIVITY.			
CC	-!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.			
CC	-!- SUBCELLULAR LOCATION: Type I membrane protein.			
CC	-!- TISSUE SPECIFICITY: HEMATOPOIETIC STEM AND PROGENITOR CELL-ENRICHED POPULATIONS. FOUND IN BRAIN, PLACENTA AND TESTIS.			
CC	-!- SIMILARITY: BELONGS TO THE CSF-1/PDGFR RECEPTOR FAMILY OF TYROSINE-PROTEIN KINASES.			
CC	-!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			

DR EMBL; M64689; AAA37634.1; -;  
 DR EMBL; X59398; CAA42041.1; -;  
 DR PIR; A39931; A39931.  
 DR PIR; S18827; S18827.  
 DR HSP; P11362; 1PGK.  
 DR MGD; MGI:95559; Flt3.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003599; Ig.  
 DR InterPro; IPR003006; Ig.MHC.  
 DR InterPro; IPR000719; Prot.kinase.  
 DR InterPro; IPR001824; RTKinaseII.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR Pfam; PF00047; Ig; 1.  
 DR Pfam; PF00069; pkinase; 1.  
 DR ProDom; PD000001; Prot.kinase; 2.  
 DR SMART; SM00409; IG; 2.  
 DR SMART; SM00219; TyKc; 1.  
 DR PROSITE; PS00835; IG-LIKE; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.  
 DR PROSITE; PS00240; RECEPTOR TYR KIN III; 1.  
 KW Signal; Transferase; Tyrosine-protein kinase; Receptor; Transmembrane;  
 KW Glycoprotein; Phosphorylation; ATP-binding; Immunoglobulin domain.  
 FT SIGNAL 1 27 POTENTIAL.  
 FT CHAIN 28 992 FL CYTOKINE RECEPTOR.  
 FT DOMAIN 28 544 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 545 564 POTENTIAL.  
 FT DOMAIN 565 992 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 254 344 IG-LIKE C2-TYPE.  
 FT DOMAIN 611 946 PROTEIN KINASE.  
 FT NP\_BIND 617 625 ATP (BY SIMILARITY).  
 FT BINDING 645 645 ATP (BY SIMILARITY).  
 FT ACT\_SITE 814 814 BY SIMILARITY.  
 FT CARBOHYD 44 44 N-LINKED (GLNAC. . .) (POTENTIAL).  
 FT CARBOHYD 133 133 N-LINKED (GLNAC. . .) (POTENTIAL).  
 FT CARBOHYD 152 152 N-LINKED (GLNAC. . .) (POTENTIAL).  
 FT CARBOHYD 307 307 N-LINKED (GLNAC. . .) (POTENTIAL).  
 FT CARBOHYD 324 324 N-LINKED (GLNAC. . .) (POTENTIAL).  
 FT CARBOHYD 352 352 N-LINKED (GLNAC. . .) (POTENTIAL).  
 FT CARBOHYD 445 445 N-LINKED (GLNAC. . .) (POTENTIAL).  
 FT CARBOHYD 474 474 N-LINKED (GLNAC. . .) (POTENTIAL).  
 FT CARBOHYD 503 503 N-LINKED (GLNAC. . .) (POTENTIAL).  
 FT CARBOHYD 542 542 N-LINKED (GLNAC. . .) (POTENTIAL).  
 FT CONFLICT 150 150 R -> A (IN REF. 2).  
 FT CONFLICT 242 242 C -> S (IN REF. 2).  
 FT CONFLICT 726 726 S -> P (IN REF. 2).  
 FT CONFLICT 957 979 CIRTSHLPKQAAPOQRGGRLAQ -> MYQNMGGNVPEHPS  
 FT CONFLICT 983 983 IYQNRPLSREAGSEPP (IN REF. 2).  
 FT CONFLICT 983 983 R -> A (IN REF. 2).  
 FT SEQUENCE 992 AA; 112639 MW; 407A087853372100 CRC64;

Query Match 100.0%; Score 5264; DB 1; Length 992;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALAQRDRRLLLVLSVMILETTVNDLPVVKVLSHENNGSSAGKPPSYRMVRGS 60  
 DB 1 MRALAQRDRRLLLVLSVMILETTVNDLPVVKVLSHENNGSSAGKPPSYRMVRGS 60  
 QY 61 PEDLQCTPRQSGTGYEATVEAEGSITLQVLAATPGDLSCLVFKHSSLGCCQPHFD 120  
 DB 61 PEDLQCTPRQSGTGYEATVEAEGSITLQVLAATPGDLSCLVFKHSSLGCCQPHFD 120  
 QY 121 LQNRGTVSMALLNVITQAGEYLLHQTOSRANYTVLTVNVRTQLYVLRPFRKQVNO 180  
 DB 121 LQNRGTVSMALLNVITQAGEYLLHQTOSRANYTVLTVNVRTQLYVLRPFRKQVNO 180  
 QY 181 DALLCISEGVPEPTVEMVLCSSHRESCKEGPAAVVRKEEKVLHELFGTDIRCCARNALGR 240  
 DB 181 DALLCISEGVPEPTVEMVLCSSHRESCKEGPAAVVRKEEKVLHELFGTDIRCCARNALGR 240  
 QY 241 ECTKLTIDLNOAPOSTLPQLFLKVGEPWIWIRCKAIHVNHGFLTWELDKALEEGSYFE 300

DB 241 ECTKLTIDLNOAPOSTLPQLFLKVGEPWIWIRCKAIHVNHGFLTWELDKALEEGSYFE 300  
 QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCSSSKHPQSALVTILEKGFINATSQEEY 360  
 DB 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCSSSKHPQSALVTILEKGFINATSQEEY 360  
 QY 361 EIDPYEKFCSFVRKAYPRIRCTWIFSQASFCEQRLGLEDGYSISKPCDHKNKPGYIFY 420  
 DB 361 EIDPYEKFCSFVRKAYPRIRCTWIFSQASFCEQRLGLEDGYSISKPCDHKNKPGYIFY 420  
 QY 421 AENDDAQFTKMTNLIRKPOVLANASASQASCSDDGYPLPSWTWKCDKSKNCTEEIP 480  
 DB 421 AENDDAQFTKMTNLIRKPOVLANASASQASCSDDGYPLPSWTWKCDKSKNCTEEIP 480  
 QY 481 EGVNKKANRKFVGOWVSSSTLNMBEAGKLLVKCCAYNSMGTSCTIFLNSGPPFPFIQ 540  
 DB 481 EGVNKKANRKFVGOWVSSSTLNMBEAGKLLVKCCAYNSMGTSCTIFLNSGPPFPFIQ 540  
 QY 541 DNISFYATIGLCLPFIVLIVLICHKKOPRYESQLOMIQVTGPLDNBYFYVDFRDY 600  
 DB 541 DNISFYATIGLCLPFIVLIVLICHKKOPRYESQLOMIQVTGPLDNBYFYVDFRDY 600  
 QY 601 DLKWEPPRENLEFGKVLGSGAGFVNMNATAYGISTGTSIQAVVMLKADSKADSKCEALM 660  
 DB 601 DLKWEPPRENLEFGKVLGSGAGFVNMNATAYGISTGTSIQAVVMLKADSKADSKCEALM 660  
 QY 661 SELKWMTHLGHNDINVLGACTLSGPVYLIFEYCCYGLLLNLYLRKREKFRHTWTEIFK 720  
 DB 661 SELKWMTHLGHNDINVLGACTLSGPVYLIFEYCCYGLLLNLYLRKREKFRHTWTEIFK 720  
 QY 721 ENHFSYPTFQAHSNSMFGSREVQLHPDLQSLGFNGNSIHSEDEIYEYENOKRLAE 780  
 DB 721 ENHFSYPTFQAHSNSMFGSREVQLHPDLQSLGFNGNSIHSEDEIYEYENOKRLAE 780  
 QY 781 EDLNVLTPEDDLCPAYQVAKMEFLFKSCVHRDLAARNVLVTHGKVKVLCDFGLARDIL 840  
 DB 781 EDLNVLTPEDDLCPAYQVAKMEFLFKSCVHRDLAARNVLVTHGKVKVLCDFGLARDIL 840  
 QY 841 SDSSVYVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900  
 DB 841 SDSSVYVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900  
 QY 901 FYKLQSGFKMEQPFVATYEGYFVWQSCWAFDSRKRKSPNLTSLFLGCOLAEAEACIRT 960  
 DB 901 FYKLQSGFKMEQPFVATYEGYFVWQSCWAFDSRKRKSPNLTSLFLGCOLAEAEACIRT 960  
 QY 961 SIHLPKQAAPOQRGGRLAQSPQVKIHRERS 992  
 DB 961 SIHLPKQAAPOQRGGRLAQSPQVKIHRERS 992

RESULT 2  
 FLT3\_HUMAN STANDARD; PRT; 993 AA.  
 ID AC P3688; Q13414;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE FL cytokine receptor precursor (EC 2.7.1.112) (Tyrosine-protein kinase  
 DE receptor FLT3) (Stem cell tyrosine kinase 1) (STK-1) (CD135 antigen).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94119906; PubMed=7507245;  
 RA Small D., Levenstein M., Kim E., Garow C., Amin S., Rockwell P.,  
 RA Witte L., Burrow C., Ratajczak M.Z., Gewirtz A.M., Civin C.I.,  
 RT "STK-1, the human homolog of Flk-2/Flt-3, is selectively expressed in  
 CD34+ human bone marrow cells and is involved in the proliferation of



early progenitor/stem cells.";  
 (2)  
 RL MEDLINE=93357464; PubMed=8394751;  
 RP Toiron O., Schiff C., Pebusque M.J., Marchetto S., Tonnelle C.,  
 RA "Human FLT3/FLK2 gene: cDNA cloning and expression in hematopoietic  
 RT cells.";  
 RL Blood 82:1110-1119(1993).  
 (3)  
 RN SEQUENCE OF 783-942 FROM N.A.  
 RP TISSUE=Testis;  
 RC MEDLINE=91169547; PubMed=2004790;  
 RA Roanet O., Mattei M.-G., Marchetto S., Birnbaum D.;  
 RT "Isolation and chromosomal localization of a novel FMS-like tyrosine  
 RL kinase gene.";  
 RL Genomics 9:380-385(1991).  
 CC -!- FUNCTION: RECEPTOR FOR THE FL CYTOKINE. HAS A TYROSINE-PROTEIN  
 CC KINASE ACTIVITY.  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- TISSUE SPECIFICITY: BONE MARROW CELLS.  
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
 CC PROTEIN KINASES.  
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
 CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD135 entry;  
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd135.htm".  
 CC -!- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;  
 CC WWW="http://www.infobiogen.fr/services/chronocancer/Genes/FLT3ID144.html".  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 DR EMBL; U02687; AAA18947.1; -;  
 DR EMBL; Z26852; CA881393.1; -;  
 DR EMBL; L36162; AAA35487.1; -;  
 DR PIR; A39061; A39061.  
 DR HSSP; P11362; 1FGK.  
 DR Genew; HGNC:3765; FLT3.  
 DR MIM; 136351; -;  
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.  
 DR GO; GO:0005021; P:vascular endothelial growth factor receptor...; TAS.  
 DR GO; GO:0008284; P:positive regulation of cell proliferation; TAS.  
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin...; TAS.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003599; Ig.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR001824; RTKinaseII.  
 DR Pfam; PF00047; Ig\_1.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Prodom; PD000001; Prot\_kinase; 2.  
 DR SMART; SM00409; IG; 1.  
 DR SMART; SM00219; TyrcK; 1.  
 DR PROSITE; PS00835; IG LIKE; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00011; PROTEIN KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN KINASE\_TYR; 1.  
 DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
 KW Signal; Transferrase; Tyrosine-protein kinase; Receptor; Transmembrane;  
 KW Glycoprotein; phosphorylation; ATP-binding; Immunoglobulin domain.  
 FT SIGNAL 1 26 POTENTIAL.  
 FT CHAIN 27 993 FL CYTOKINE RECEPTOR.  
 FT DOMAIN 27 543 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 544 563 POTENTIAL.

FT	DOMAIN	564	993	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	253	343	IG-LIKE C2-TYPE.
FT	DOMAIN	610	943	PROTEIN KINASE.
FT	NP_BIND	616	624	ATP (BY SIMILARITY).
FT	BINDING	644	644	ATP (BY SIMILARITY).
FT	ACT_SITE	811	811	BY SIMILARITY.
FT	CARBOHYD	43	43	N-LINKED (GLCNAC...)
FT	CARBOHYD	100	100	N-LINKED (GLCNAC...)
FT	CARBOHYD	151	151	N-LINKED (GLCNAC...)
FT	CARBOHYD	306	306	N-LINKED (GLCNAC...)
FT	CARBOHYD	323	323	N-LINKED (GLCNAC...)
FT	CARBOHYD	351	351	N-LINKED (GLCNAC...)
FT	CARBOHYD	354	354	N-LINKED (GLCNAC...)
FT	CARBOHYD	473	473	N-LINKED (GLCNAC...)
FT	CARBOHYD	502	502	N-LINKED (GLCNAC...)
FT	CARBOHYD	541	541	N-LINKED (GLCNAC...)
FT	CONFLICT	8	8	A -> G (IN REF. 2).
FT	CONFLICT	10	11	TV -> QL (IN REF. 2).
FT	CONFLICT	78	78	A -> R (IN REF. 2).
FT	CONFLICT	227	227	T -> M (IN REF. 2).
FT	CONFLICT	346	346	G -> E (IN REF. 2).
FT	CONFLICT	940	940	T -> H (IN REF. 3).
SQ	SEQUENCE	993 AA;	112804 MW;	16790124B02F6BBF CRC64;

Query Match 84.1%; Score 4429.5; DB 1; Length 993;  
 Best Local Similarity 84.1%; Pred. No. 9e-290; Indels 11; Gaps 4;  
 Matches 836; Conservative 57; Mismatches 90;

QY	1	MRLAQRSDRLLLVLVLVLSMILEVTNQDLPIVKCVLISHENNGSSAGKPSRYMRVGS	60
DB	1	MPALA-RDAGTVPLLVVFSAMIFGTITQDLPIVKCVLHNKNDSSVGKSSYPNVSES	59
QY	61	PEDLOCTPRRSEGTVEAATVEVAESGSIITLQVLATPGDLSCLWVFKHSLGCOPHPD	120
DB	60	PEDLGALRPQSGTVEAAAEVDVSASITLQVLDAFGNISCLWVFKHSLINCQHPD	119
QY	121	LQNRGIVSMALINVTETQAGEVLLHIQSERANYTVLFTVNVDRDTQLVLRFRYFRKQENQ	180
DB	120	LQNRGVSMVLKMTETQAGEVLLFIQSEATNYTILFTVSIKNTLLYTLRRYFRKQENQ	179
QY	181	DALLCISEGVPEPTVEWVLCSSHRESSCKEESGPAVVRKEEKLHFLFGTDIRCARNALGR	240
DB	180	DALVCISESVPEPIVEWVLCDSQGESCKEESPAVVRKEEKLHFLFGTDIRCARNELGR	239
QY	241	ECKLFTIDLNAQPOSTLPQLKYGEPLWIRCKAIHVNHGFGLTWELEDKALEGSYPE	300
DB	240	ECKLFTIDLNAQPOSTLPQLKYGEPLWIRCKAVHVNHGFGLTWELENKALEGSYPE	299
QY	301	MSTYSTNTRMIRILLAFVSSVGRNDTGYYTSSSKHPQSQSALVTILEKGFINATSSQEEY	360
DB	300	MSTYSTNTRMIRILLAFVSSVARNDTGYYTSSSKHPQSQSALVTIVGKGFINATSSSEY	359
QY	361	EIDPYEKEFCFSVRKAYPRICTWIFSOASPFCEORGLDGYISIKFCDHKNKPGYIFY	420
DB	360	EIDQYEEFCFSVRKAYPQIRCTWTFSRKSPFCEORGLDNGYSISKFCNHRKQPGYIFYH	419
QY	421	AENDDAQFTKMTLNIKKKPOVLANASQASCSGSDGYPLPSWTWKCKSDKSPNCTEEIP	480
DB	420	AENDDAQFTKMTLNIKKKPOVLAEASQASCSGSDGYPLPSWTWKCKSDKSPNCTEEIT	479
QY	481	EGVWNKANKRVFGQWVSSSTLNMSAGKLLVKKCAVNSMGTSCETIFLNSPGPPFFIQ	540
DB	480	EGVWNKANKRVFGQWVSSSTLNMSAINGFLVKCAVNSLGTSCETILLNSPGPPFFIQ	539
QY	541	DNISFYATIGLCLPIVLIIVLICHYKQKPYESQLOMIQVTPGLDNEFYVYDPRDYEY	600
DB	540	DNISFYATIGVCLLFIIVLTLCHYKQKPYESQLOMQVQVGTGSSDNEIFYVYDFREY	599
QY	601	DLKWEFFRENLEFGKVLGSGAFGRVNNATAYGISKTGVSIQVAVKMLKEKADSCKEALM	660
DB	600	DLKWEFFRENLEFGKVLGSGAFGRVNNATAYGISKTGVSIQVAVKMLKEKAUSEREALM	659
QY	661	SELKVMTHLGHHDNIVNLLGACTLSGVPVLIPEYCCYGDLLNLYLRSKREKFRHTWTETPK	720

```

Db 660 SELKMTQLGSHENIVNLLGACTLSGPIYLIFEYCCYGDLLNLRKREKPHRTWTEIPK 719
Qy 721 EHNFSYPTFOAHSSNMPGSREVOLHPPDQLSGFNGSHSEDEIYEENOKRLAEEB 780
Db 720 EHNFSYPTFOAHSSNMPGSREVOLHPPDQLSGFNGSHSEDEIYEENOKRLAEEB 777
Qy 781 EDNLVLTFFDLFCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840
Db 778 EDNLVLTFFDLFCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIM 837
Qy 841 SDSSVVRGNARLPVKWAPESLFEIGIYIKSDVMSYGILLWEIFSLGWNYPGIPVDAN 900
Db 838 SDSSVVRGNARLPVKWAPESLFEIGIYIKSDVMSYGILLWEIFSLGWNYPGIPVDAN 897
Qy 901 FYKLQSGFKMQPFVATEGIYFVMSQWAFPSRKRPSFNLTSFLGCOLAABEAC--- 957
Db 898 FYKLQSGFKMQPFVATEGIYFVMSQWAFPSRKRPSFNLTSFLGCOLAABEAMWON 957
Qy 958 ----IRTSIHLPKQAPOORG-GLRAQSPQOVK 986
Db 958 VDRGVSECPHTYQNRFPSPREMDLGLLSPOAQVE 991

RESULT 3
KFMS FELCA
ID_KFMS FELCA STANDARD; PRT; 980 AA.
AC P13369;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Macrophage colony stimulating factor I receptor precursor (CSF-1-R)
DE (EC 2.7.1.112) (fms proto-oncogene) (c-fms).
GN CSF1R OR FMS.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89077553; PubMed=2849512;
RA Woolford J., McAuliffe A., Rohrschneider L.R.;
RT "Activation of the feline c-fms proto-oncogene: multiple alterations
RT are required to generate a fully transformed phenotype.";
RL Cell 55:965-977(1988).
CC -!- FUNCTION: THIS PROTEIN IS THE RECEPTOR FOR CSF-1, IT IS A PROTEIN
CC -!- TYROSINE-KINASE TRANSMEMBRANE RECEPTOR.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC -!- tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC -!- PROTEIN KINASES.
CC -!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J03149; AAA30811.1; -.
CC PIR; A31636; TVCTMD.
CC HSP; P11362; IFKC.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003598; Ig C2.
CC InterPro; IPR003006; Ig MHC.
CC InterPro; IPR000719; Prot kinase.
CC InterPro; IPR001824; RTkinaseII.
CC InterPro; IPR001245; Tyr_pkinase.
CC Pfam; PF00047; Ig; 3.
CC Pfam; PF00069; pkinase; 1.

```

```

DR ProDom; PD000001; Prot_kinase; 2.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00219; TyrKC; 1.
DR PROSITE; PS00835; IG LIKE; 3.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
KW Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
KW Immunoglobulin domain; Repeat
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 980 MACROPHAGE COLONY STIMULATING FACTOR I
FT RECEPTOR.
FT EXTRACELLULAR (POTENTIAL).
FT DOMAIN 20 509 POTENTIAL.
FT TRANSMEM 510 535 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 536 980
FT DOMAIN 21 100 IG-LIKE C2-TYPE 1.
FT DOMAIN 107 197 IG-LIKE C2-TYPE 2.
FT DOMAIN 202 297 IG-LIKE C2-TYPE 3.
FT DOMAIN 299 397 IG-LIKE C2-TYPE 4.
FT DOMAIN 400 499 IG-LIKE C2-TYPE 5.
FT DOMAIN 579 908 PROTEIN KINASE
FT NP_BIND 585 593 ATP (BY SIMILARITY).
FT BINDING 613 613 ATP (BY SIMILARITY).
FT ACT_SITE 776 776 BY SIMILARITY.
FT DISULFID 42 84 POTENTIAL.
FT DISULFID 127 177 POTENTIAL.
FT DISULFID 224 278 POTENTIAL.
FT DISULFID 417 482 POTENTIAL.
FT MOD_RES 807 807 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CARBOHYD 45 45 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 73 73 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 94 94 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 153 153 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 275 275 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 286 286 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 302 302 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 335 335 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 410 410 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 477 477 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 490 490 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 980 AA; 108506 MW; 4B5CF6E1B97CF7CFCF CRC64;

Query Match 23.9%; Score 1258.5; DB 1; Length 980;
Best Local Similarity 32.7%; Pred. No. 8.9e-77;
Matches 328; Conservative 168; Mismatches 353; Indels 153; Gaps 32;

Qy 75 TYVEAATVEVAESGITLQVQLATPGDLISCL-----W---VFKHSSLCQPHFDLQNRG 125
Db 13 TAWHAQGVPIQSPGPELVVEPGTTVLRVCGVSGVEMDGPISPHWNLDLDPSSI---- 68

Qy 126 IVSMAILNVTQAGEYLLHI-----QSERANTVTLFTVNRDTQLYVL---RRPY----- 173
Db 69 ---LTTNNATFQNTGY--HCTEPGNPOGGNATI-----HLVYKDPARPKVLAOE 114

Qy 174 FRKMENODALL-CISEGPEPTVEVLCSHRESCKEKGPAVVRKEEVKVLHELFGTDI-- 230
Db 115 VTVLEGQDALFPLC---LTDPALE-----AGSVLRVRGRPVLRQTNVTSFSPMHGFTIHK 166

Qy 231 -----RCCARNALGRECTKL-----FTIDLQAPOSTL--POLFLKV-GEPLWIRCK 274
Db 167 AKFIENHYVQCSAR--VDGRVTVMGILWKVOKDISGPATLTLEPAELVIOGEAAQIVCS 225

Qy 275 AIHVNHGFLTWELDKALEBGSYFEMSTYSTNRNTRIRILLAFVSVSGRNDTGYTCSSS 334
Db 226 ASNIDWNFDVSLRHGDTKL---TISQQSDFDHNRVQ-KVLTNLNLDHVSFDAGNYSCTAT 281

Qy 335 K---HPQSALVTILEKGFINATSSOE-EYEDLPYEKCFCSYRFAKAYPRIRC-TWI---- 385
Db 282 NAWGNHSAVMFRVVSAYLNLTSQSULLQEVTVGKVDLQVKVAYPGLESFNWYTLGP 341

Qy 386 FSQASFPCEORGLDGYSTSKFCD-----HKNKPGVEYIFYAENDDAQFTQKFTLNIRKKPO 441

```

Db 342 FSDYQDKDFVIKTYRTYRTSLPRLKRSSEAGRYSLARAGGQNALTFELTLRYPE 401  
QY 442 VLANASAOAS---CSSDGYPLPWTWKCKSDKSPNCTEE---IPEGVWNKANKRVFG 494  
Db 402 VRVVTTLINGSDTLLCEASGYPOPSVTWQCRSHSTRDCDESAGLVLEDHSHSEVLQSVPFH 461  
QY 495 QWSSSTLNMBSAGKLLVCCAYNSMGTSCTIFLNSPGPPFFODNISFYATIGLC-- 552  
Db 462 EVIHSLLAIGTLEHNRNRYECRAFNVSNGSSQTFPISIGANTQLPDELFTFVLLTQMS 521  
QY 553 -LPFTWLVILVILCHYKKGFRYESOLOMTQVTPGLDNEYFYDFRDYEDLKWEPREN 611  
Db 522 IWALLLLLLLLLYKKPKQPKQVRWKIE--SYEGNSVTFIDPTQLPYNKWEPRN 579  
QY 612 EFGKVLGSAFGRVNMATYAGISKTGVSIOQAVKMLKEKADSEKALMSELKMTHTLGH 671  
Db 580 QFGKTLGAGAFKGVVEATAFGLGKEDAVLKVAVMKLTSTAHADKEALMSELKMTHTLGH 639  
QY 672 HDNIVNLGACTSGRVYLIFFYCCVGDLLNVLRSKE-----KF 711  
Db 640 HENIVNLGACTHGGFVLVITEYCCVGDLLNVLRSKE-----KF 699  
QY 712 HRTWTEIFKEHNFSSVPTFQAHSNMPGSRVQLHPDLDLQSGFNGNSIHSEDEIEYEN 771  
Db 700 HLEKYYVRSDFS-----SQGVDTYVEMRPVSTSSNDSFSEEDL---- 740  
QY 772 QKRLAEEBEDNLVLTFFEDLLCFAYQVAKMGFLFEPKSCVHRDLAARNVLVTHGKVYKIC 831  
Db 741 -----GKEDGRPLELRLDLHFFSSQVAGMAFLASKNCIHRDVAARNVLLTSGRVAKIG 793  
QY 832 DFLGARDILSDSVYVGNARLPVKMAPESIFEGYITTKSDVMSYIGILLWIFSLGVNP 891  
Db 794 DFLGARDIMDSNVIKGNARLPVKMAPESIFDCVTVQSDVMSYIGILLWIFSLGLNP 853  
QY 892 YGIPVDANFYKLIQSGFQWEPFATYEGYIFWQSCWAFDSKRPSPFNLTSLFLGCOLA 951  
Db 854 YPGILVNSFYKLVKDGQYMAQPAFAPKNIYSIMQACWALEPRTTFOQICSL--OKO 911  
QY 952 EAEACIRTSIHLKQAAPQ-----RGLRAQSPQRQ 984  
Db 912 AQEDRVRPNYTNLPSS 953  
RESULT 4  
KFMS FSVMD  
ID KFMS FSVMD STANDARD; PRT; 978 AA.  
AC P00345; Q86597; 978 AA.  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Tyrosine-protein kinase transforming protein fms (EC 2.7.1.112).  
GN V-FMS.  
OS Feline sarcoma virus (strain McDonough).  
OC Viruses; Retroviral viruses; Retroviridae; Mammalian type C retroviruses.  
OX NCBI\_TaxID=11778;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=84119469; PubMed=6582485;  
RA Hampe A., Gobet M., Sherr C.J., Galibert F.;  
RT "Nucleotide sequence of the feline retroviral oncogene v-fms shows  
RT unexpected homology with oncogenes encoding tyrosine-specific protein  
RT kinases";  
RL Proc. Natl. Acad. Sci. U.S.A. 81:85-89 (1984).  
RN [2]  
RP REVISIONS, SEQUENCE FROM N.A.  
RX MEDLINE=92015516; PubMed=1833563;  
RA Smola U., Hennig D., Hadwiger-fangmeier A., Schuetz B., Pfaff E.,  
RA Niemann H., Tamura T.;  
RT "Reassessment of the v-fms sequence: threonine phosphorylation of the  
RT COOH-terminal domain";  
RL J. Virol. 65:6181-6187 (1991).  
CC -!- FUNCTION: V-FMS IS DERIVED FROM THE RECEPTOR FOR COLONY

CC STIMULATING FACTOR 1 (CSF-1).  
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
CC tyrosine phosphate.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-FMS  
CC POLYPROTEIN.  
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
CC PROTEIN KINASES.  
CC -!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC  
CC EMBL: K01643; AAA43045.1; -.  
CC EMBL: S59588; AAB20028.1; -.  
CC PIR: A00654; TVMWD.  
CC HSP: P11362; IFCK.  
CC InterPro: IPR007110; IG-like.  
CC InterPro: IPR003598; IG C2.  
CC InterPro: IPR003006; IG\_MHC.  
CC InterPro: IPR000719; Prot\_kinase.  
CC InterPro: IPR001824; RtkinaseII.  
CC InterPro: IPR001245; Tyr\_kinase.  
CC Pfam: PF00047; ig; 3.  
CC Pfam: PF00069; pkinase; 1.  
CC ProDom: PD000001; Prot\_kinase; 2.  
CC SMART: SM00408; IGc2; 1.  
CC SMART: SM00219; Tyrc; 1.  
CC PROSITE: PS00835; IG LIKE; 3.  
CC PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
CC PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
CC PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
CC PROSITE: PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
CC PolyProtein: Tyrosine-protein kinase; Oncogene; Transferase; Receptor;  
KW Transmembrane; Glycoprotein; Phosphorylation; ATP-binding;  
KW Immunoglobulin domain; Repeat.  
FT DOMAIN 1 543 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 544 588 POTENTIAL.  
FT DOMAIN 569 978 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 55 134 IG-LIKE C2-TYPE 1.  
FT DOMAIN 141 231 IG-LIKE C2-TYPE 2.  
FT DOMAIN 236 331 IG-LIKE C2-TYPE 3.  
FT DOMAIN 333 431 IG-LIKE C2-TYPE 4.  
FT DOMAIN 434 533 IG-LIKE C2-TYPE 5.  
FT DOMAIN 613 942 PROTEIN KINASE.  
FT NP\_BIND 619 627 ATP (BY SIMILARITY).  
FT BINDING 647 647 ATP (BY SIMILARITY).  
FT DISULFID 76 118 POTENTIAL.  
FT DISULFID 161 211 POTENTIAL.  
FT DISULFID 258 312 POTENTIAL.  
FT DISULFID 451 516 POTENTIAL.  
FT MOD\_RES 841 841 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
FT MOD\_RES 973 973 PHOSPHORYLATION.  
FT CARBOHYD 79 79 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 107 107 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 128 128 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 187 187 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 309 309 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 320 320 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 369 369 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 444 444 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 511 511 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 524 524 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CONFLICT 714 714 L -> P (IN REF. 1).  
FT CONFLICT 971 978 ORTPPVAR -> RGPL (IN REF. 1).  
FT SEQUENCE 978 AA; 108491 MW; 4C7CAC4835185EBF CRC64;

```
Query Match      23.7%; Score 1249.5; DB 1; Length 978;
Best Local Similarity 33.2%; Pred. No. 3.6e-76;
Matches 328; Conservative 162; Mismatches 348; Indels 149; Gaps 32;

Qy 75 TYVEAATVEVAGSGSTITLOVQATPGDLSCL-----W---VPKHSLSGCPHFQDLQNRG 125
Db 47 TAWHAQGVPIQSPGPELVFPGTTVTRCVGNGSVENDGPISPHWNLDLPSSI-----102
Qy 126 IVSMALNVETQAGGYLLHI-----QSERANVTFLTVNRDQLVYL---RRPY-----173
Db 103 ---LTTNATFQNTGY---HCTEPGNPRGNATI-----HLVXDPAPKPKVLAQE 148
Qy 174 FRKMENQDALL-CISGVPEPTVEWVLCSSHRESCKEKGPAVVRKEKVLHFLGTDI--230
Db 149 VTVLEQDALLPCL---LTDPALE-----AGSLVRVRGRPLVQNTNYSFSPWHGFTIHK 200
Qy 231 -----RCARNALGRECTKL-----FTIDLQAQOSTL-PQLFLKV-GEPLWIRCK 274
Db 201 AKFIENHVYOCGAR-VDRGTVTSMGLWLVKQKDIDGSPATLTLEPAELVRIQGEAAQIVCS 259
Qy 275 AIHVNHGFLTWELEDKALEEGSYFMSTYTNRTMIRILLAFVSVGRNDTGYYTCSSS 334
Db 260 ASNIDVNFVLSRHGDTKL---TISQSDPHDNRYQ-KVLTNLNDHVSFDQAGNYSCTAT 315
Qy 335 K---HPSQALVTILEKGFINATSSQE-EYEIDPYEKFCFSVRFKAYPRIRC-TWI-----385
Db 316 NAWGNHSAWMPRVVBSAYSNLTSQSLQEVTVGEKVDLVQVKEAYPGLESFNMWYLG 375
Qy 386 FQASFPFCQRGLEDGYSISKCD-----HKNRGEYIFYAENDDAQFTKMTLINIRKXQ 441
Db 376 FSDYQDKLDFVTKIDRYTSTLSLRLKRSESGRYSLARNAGGQNALTFELTYRYPE 435
Qy 442 VLANASASQAS-----CSSDQPLPSTWKKCSKSPNCTEE-----IPEGVWKNKANKKVF 494
Db 436 VRVTWTLINGSDTLLCEASGYPOPSVTWVQCRSHTDRCDSEAGLVDSHSEVLSQVPY 495
Qy 495 QWVSSSTLNMSEAGKLLVKKCAYNMGTSCTETIFLNSPFPFFIQDNISFYATIGLC--552
Db 496 EVIVHSLLAIGTLEHNRVTCEAFNVSNGSSQTFWPISIGATPLPDELLFTVPLVLTCS 555
Qy 553 -LPFIVVLIVLCHKYKKQFRYESQLOMTQVGTPLDNEYFYDFRDEYDLKWEFFREN 611
Db 556 IMALLLLLLLYLYKQPKYQVRWKIIE--SYEGNSYTFIDPTQLPYNEKWEFFRN 613
Qy 612 BFGKVLGSGAFGRVMNATAVGSKTGVTSQVAVMKLKERADSCKEKALMSELKQMTLGH 671
Db 614 QFGKLTGTGAFKGVKVEATAPGLGKEDAVLVKAVVMLKSTAHADKELMSELKIMHLGQ 673
Qy 672 HDNIVNLLGACTLSGFPVLIIFYCCYGDLLNLYLSKRE-----KF 711
Db 674 HENIVNLLGACTHGGFPVLVITEYCCYGDLLNLFRRQAEAMLGPSLVGQDPEAGAGYKNI 733
Qy 712 HRTWTEIFKEHNPSSYPTFOAHNSNMPGSRREVQLHPDLQSGFNGNSIHSEDEIYEN 771
Db 734 HLEKKYVRDSDGFS-----SQGVDTYVEMRPVSTSSNDSFSEEDL-----774
Qy 772 QKRLAEEBEDNLVTFEDLFCAYOVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKIC 831
Db 775 -----GKEDGRPLERDLHSSQVQAQMAFLSKNCIHRDVAARNVLLTSGRVAKIG 827
Qy 832 DFGLARDILSDSYVYVGNARLPVKWMAPESLFEGIYTTIKSDVMSYGILLWEIFSLGVNP 891
Db 828 DFGLARDINDNSYIVKGNARLPVKWMAPESIDFCVTVQSDVMSYGILLWEIFSLGLNP 887
Qy 892 YGIPVDANFYKLIQSGFKWQBPFYATEGIYFWQSCWAPDSKRKRSFPNLTFLGCOLA 951
Db 888 YFGILVNSFYKLVKDGQYQMAQAPAFAPKNIYSIMQACWALEPRTTRPTFQICSL--QK 945
Qy 952 EAEACIRTSIHLPKQAA-----PQOR 973
Db 946 AOEDRRVPNTNLLPSSSSSRLRPWQR 972
```

```
RESULT 5
KIT MOUSE
ID KIT MOUSE STANDARD; PRT; 975 AA.
AC P05532; O61415; O61416; O61417;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Mast/stem cell growth factor precursor (EC 2.7.1.112) (SCFR)
DE (Proto-oncogene tyrosine-protein kinase kit) (C-kit).
GN KIT OR SL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE=88296403; PubMed=2456920;
RA Qiu F., Ray P., Brown K., Barker P.E., Jhanwar S., Ruddle F.H.,
RA Besmer P.;
RT "Primary structure of c-kit: relationship with the CSF-1/PDGF
RT receptor kinase family -- oncogenic activation of v-kit involves
RT deletion of extracellular domain and C-terminus.";
RL EMBO J. 7:1003-1011(1988).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM TRUNCATED).
RC STRAIN=ICR;
RX MEDLINE=92331813; PubMed=1378413;
RA Rossi P., Marziali G., Albanesi C., Charlesworth A., Gheremia R.,
RA Sorrentino V.;
RT "A novel c-kit transcript, potentially encoding a truncated receptor,
RT originates within a kit gene intron in mouse spermatids.";
RL Dev. Biol. 152:203-207(1992).
RN [3]
RP LIGAND.
RX MEDLINE=91006023; PubMed=1698611;
RA Tan J.C., Buck J., Levi E., Besmer P.;
RT "Candidate ligand for the c-kit transmembrane kinase receptor: KL, a
RT fibroblast derived growth factor stimulates mast cells and erythroid
RT progenitors.";
RL EMBO J. 9:3287-3294(1990).
RN [4]
RP VARIANT W42 ASN-790.
RX MEDLINE=90100577; PubMed=1688471;
RA Tan J.C., Nocka K., Ray P., Traktman P., Besmer P.;
RT "The dominant W42 spotting phenotype results from a missense mutation
RT in the c-kit receptor kinase.";
RL Science 247:209-212(1990).
RN [5]
RP VARIANTS W37 LYS-582; WV MET-660 AND W41 MET-831.
RX MEDLINE=90269214; PubMed=169331;
RA Nocka K., Tan J.C., Chiu E., Chu T.Y., Ray P., Traktman P.,
RA Besmer P.;
RT "Molecular bases of dominant negative and loss of function mutations
RT at the murine c-kit/white spotting locus: W37, WV, W41 and W.";
RL EMBO J. 9:1805-1813(1990).
CC -1- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
CC 3-KINASE (PI3K).
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- DISEASE: WHITE-SPOTTING VARIANTS INDUCES SEVERE EFFECTS ON
CC PIGMENTATION, GAMETOGENESIS AND HEMATOPOIESIS. MICE HOMOZYGOUS
CC FOR W42 DIE PERINATALLY OF MACROCYTIC ANEMIA.
CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -1- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```







```

RESULT 7
KFMS HUMAN
ID KFMS_HUMAN STANDARD; PRT; 972 AA.
AC P07333;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Macrophage colony stimulating factor I receptor precursor (CSF-1-R)
DE (EC 2.7.1.112) (Fms proto-oncogene) (c-fms) (CD115 antigen).
GN CSF1R OR FMS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89239490; PubMed=2524025;
RA Hampe A., Shamoon B.M., Gobet M., Sherr C.J., Galibert F.;
RT "Nucleotide sequence and structural organization of the human FMS
RT proto-oncogene.";
RL Oncogene Res. 4:9-17(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86175013; PubMed=2421165;
RA Cousens L., van Beveren C., Smith D., Chen E., Mitchell R.L.,
RA Isacke C.M., Verma I.M., Ullrich A.;
RT "Structural alteration of viral homologue of receptor proto-oncogene
RT fms at carboxyl terminus.";
RL Nature 320:277-280(1986).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=9717923; PubMed=9027509;
RA Andre C., Hampe A., Lachaume P., Martin E., Wang X.P., Manus V.,
RA Hu W.X., Galibert F.;
RT "Sequence analysis of two genomic regions containing the KIT and the
RT FMS receptor tyrosine kinase genes.";
RL Genomics 39:216-226(1997).
RN [4]
RP SEQUENCE OF 1-16 FROM N.A.
RX MEDLINE=89261741; PubMed=2524648;
RA Visvader J., Verma I.M.;
RT "Differential transcription of exon 1 of the human c-fms gene in
RT placental trophoblasts and monocytes.";
RL Mol. Cell. Biol. 9:1336-1341(1989).
RN [5]
RP SEQUENCE OF 1-16 FROM N.A.
RX MEDLINE=86218020; PubMed=3525854;
RA Wheeler E.F., Roussel M.F., Hampe A., Walker M.H., Fried V.A.,
RA Look A.T., Rettenmier C.W., Sherr C.J.;
RT "The amino-terminal domain of the v-fms oncogene product includes a
RT functional signal peptide that directs synthesis of a transforming
RT glycoprotein in the absence of feline leukemia virus gag sequences.";
RL J. Virol. 59:224-233(1986).
CC -!- FUNCTION: THIS PROTEIN IS THE RECEPTOR FOR CSF-1, IT IS A PROTEIN
CC -!- TYROSINE-KINASE TRANSMEMBRANE RECEPTOR.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane receptor.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
CC -!- DATABASE: NAME=PROW; NOTE=PROW 1:13-21(2000);
CC WWW="http://www.ncbi.nlm.nih.gov/prov/guide/1999017666.g.htm".
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
Query Match 23.2%; Score 1221.5; DB 1; Length 972;
Best Local Similarity 32.5%; Pred. No. 2.7e-74;
Matches 330; Conservative 152; Mismatches 331; Indels 201; Gaps 33;

```







QY 927 SCHAFDSRKRPSFPNLTSLFCQLAE 953  
 Db 906 TCWDADPLKRPTEFKQIVQLIEKQISES 932

RESULT 10  
 KPM5 RAT  
 ID KPM5 RAT STANDARD; PRT; 978 AA.  
 AC Q00495;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Macrophage colony stimulating factor I receptor precursor (CSF-1-R)  
 DE (EC 2.7.1.112) (Pms proto-oncogene) (c-fms).  
 GN CSF1R OR CSF1R OR FMS.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Wistar; TISSUE=Skeletal muscle;  
 RX MEDLINE=93001225; PubMed=139227;  
 RA Borycki A.G., Guiller M., Leibovitch M.P., Leibovitch S.A.;  
 RT "Molecular cloning of CSF-1 receptor from rat myoblasts. Sequence  
 analysis and regulation during myogenesis.";  
 RL Growth Factors 6:209-218(1992).  
 CC -!- FUNCTION: THIS PROTEIN IS THE RECEPTOR FOR CSF-1, IT IS A PROTEIN  
 CC TYROSINE-KINASE TRANSMEMBRANE RECEPTOR.  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
 CC PROTEIN KINASES.  
 CC -!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))

EMBL; X61479; CAA43706.1; --  
 PIR; I60321; S16385.  
 HSP; P11362; IFCG.  
 InterPro; IPR007110; Ig-like.  
 InterPro; IPR003598; Ig C2.  
 InterPro; IPR003006; Ig MHC.  
 InterPro; IPR000719; Prot kinase.  
 InterPro; IPR001824; TKinaseIII.  
 InterPro; IPR001245; Tyr\_kinase.  
 Pfam; PF00047; Ig; 4.  
 Pfam; PF00069; pkinase; 1.  
 ProDom; PD000001; Prot kinase; 2.  
 SMART; SM00408; IGC2; 1.  
 SMART; SM00219; TyKc; 1.  
 PROSITE; PS00835; Ig-Like; 4.  
 PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 PROSITE; PS50011; PROTEIN KINASE DOM; 1.  
 PROSITE; PS00109; PROTEIN KINASE TYR; 1.  
 PROSITE; PS00240; RECEPTOR TYR KIN III; 1.  
 KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;  
 KW Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;  
 KW Immunoglobulin domain; Repeat.  
 SIGNAL 1 19  
 CHAIN 20 978  
 MACROPHAGE COLONY STIMULATING FACTOR I  
 RECEPTOR.  
 DOMAIN 20 511  
 EXTRACELLULAR (POTENTIAL).  
 TRANSMEM 512 536  
 POTENTIAL.  
 DOMAIN 537 978  
 CYTOPLASMIC (POTENTIAL).  
 DOMAIN 24 104  
 IG-LIKE C2-TYPE 1.

FT DOMAIN 107 197 IG-LIKE C2-TYPE 2.  
 FT DOMAIN 204 298 IG-LIKE C2-TYPE 3.  
 FT DOMAIN 299 397 IG-LIKE C2-TYPE 4.  
 FT DOMAIN 398 503 IG-LIKE C2-TYPE 5.  
 FT DOMAIN 580 914 PROTEIN KINASE.  
 FT NP BIND 586 594 ATP (BY SIMILARITY).  
 FT BINDING 614 614 ATP (BY SIMILARITY).  
 FT ACT SITE 776 776 BY SIMILARITY.  
 FT DISULFID 42 84 POTENTIAL.  
 FT DISULFID 127 177 POTENTIAL.  
 FT DISULFID 224 278 POTENTIAL.  
 FT DISULFID 417 483 POTENTIAL.  
 FT MOD RES 706 706 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT CARBOHYD 45 45 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 73 73 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 302 302 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 335 335 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 389 389 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 410 410 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 449 449 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 478 478 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 491 491 N-LINKED (GLCNAC...) (POTENTIAL).  
 SQ SEQUENCE 978 AA; 109264 MW; 0A68456EF56BC7E3 CRC64;  
 Query Match 23.0%; Score 1210; DB 1; Length 978;  
 Best Local Similarity 33.5%; Pred. No. 1.6e-73;  
 Matches 318; Conservative 147; Mismatches 329; Indels 154; Gaps 31;  
 QY 78 EAATVEAEGSGITLQVQLATPGDLSCLMVKHSSGCGPHFDLQNGRIVSMALNVET 137  
 Db 36 ETVLRCSVNGSVEWD-----GPISPYWTLDPESPGS-----TLTRNATPK 77  
 QY 138 QAGEY-LLHIQSERANYTVLPTVNVDTOLYV-----LRRPYFRKMNQDALL-CIS 187  
 Db 78 NTGYTRCTELEDPMAGSTI-----HLVYKDPAHSNWLLAQEVTVVGEQEAVALPCL- 128  
 QY 188 EGVPEPTVEVWLCSHRESCKEEGPAVVRK-----EBKVLHELFGTDIRCC 233  
 Db 129 -ITDPALK-----DSVSLMRGGQVLKRTVYFSAWRGFIRKAKVL-----DSNTYVC 177  
 QY 234 ARNALGRECTKL-FTIDLN---QAPQSTL-PQLFLKV-GEPLWIRCKAIHVNHGFLTW 286  
 Db 178 KTWNGREESTGTGLKVNVRVHPEPPQIKLEPSPKLVIRGEAAQIVCSATNAEVEGNVL 237  
 QY 287 ELEDKALE--EGSYFEMSTYSTNRTMIRLLAFVSSVGRNDTGYTC-SSSKHPQSALV 343  
 Db 238 KRGDTKLEIPLNSDFQDNYIKKVRAL-----SLNAVDFQDAGIYSCVASNDVGRTRATM 291  
 QY 344 T--ILEKGFINATSSQE--EYEDIDPVKFCFVRPKAYPRIR-CTWIFQASPPCQORGL 399  
 Db 292 NFQVVESAYLNLTSQSLQEVSVGDSLLITVHADAYPSIQHYNWYTLGPPFE--DORKLE 350  
 QY 400 -----DGYSISKCDH--KMKPEGYIFYAENDDAQFTKMTLIRKPKOVLANASAQ 450  
 Db 351 FITQRIAYRYTFKLFLNRVRKASEAGQYFLMAQKAGWNNLTFELTLRYPPEVSVTWMPVN 410  
 QY 451 AS---CSSDGYPLPSWTWKCKSDKSPNCTEIEPEGVWN-----KKAHKVFGQWVSSSTL 502  
 Db 411 GSDVLFCDVSGYGPQPSVTWMECRGHTDRCDQAQALQVWMDTHPEVLSPKDPKVIQSQL 470  
 QY 503 NMSAAGKGLLVKCAVNSMGTSCBTIFLNSPGPPFFIQDNISFYATIGLC---LPFIIVL 559  
 Db 471 PIGTLKHNMTYFCKTHNSVGNSSQYFVAVSLGSKQKLPDESFLTTPVAVVACMSVMSLLVL 530  
 QY 560 IVLICHYKQFVBSQLOMIQ-VTGELDNEYFYVDFRDEYEDLKEFFENLEFGKVLG 618  
 Db 531 LLLLLYKQKPKYQVRWKIIIEYEG---NSYTFIDPTQLPYNEKWEFFRNLOFGKTLG 587  
 QY 619 SGAFGRVMNATYGIKSTGVSIQVAVKMLKEKADSCKEALMSLKMTHLGHNDIVNL 678  
 Db 588 AGAFGKVETAFLGKGDVAVLKVAVKMLKSTAHADKEALMSELKIMSHLGHQENIVNL 647  
 QY 679 LGACTLSGPVYLIFEYCCYGDLLNLYRSKRE-----KFRHTWTEI 718

```

648 LGACTHGGPVLVITEYCCYGDLLNFLRKAEMLGPSLPGQDSEGDSSYKNHLEKYY 707
719 FKEHNFSSYTFQAHSSNMPGSEVQLHPDLQDLSGFNGNSIHSEDEIEYENQKRLAAE 778
708 RDSGFS-----SGVDYVEMRPVSTSSDSFFKQD-----L 740
779 EEDLNVLTFEDLLCFAYQVAKNGMEFLFKSVHRDLAARNVLVTHGKWKICDFGLARD 838
741 DKSPSRLELWDLHFSQVQAQNAFLASKNCIHRDVAARNVLTSGHVAKIGDFGLARD 800
839 ILSDSSYVVRGNARLPVKWMAPELSFEGITYIKSDVMSYSGILLWEIFSLGWNPPGIPVD 898
801 IMDSNVVYKGNARLPVKWMAPELSILYCVTVQSDVMSYSGILLWEIFSLGWNPPGILVN 860
899 ANFKLQSGFKMOPFYATYEGIFYVMSQWAFDSRKPSPNLTSEL 946
861 NKFKLVKDGQMAQPVFAPKNYIYIMQSCWDLPTFRTRPTQQICFLL 908

RESULT 11
KIT_CANFA
ID - KIT_CANFA STANDARD; PRT; 975 AA.
AC O97799;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Mast/stem cell growth factor receptor precursor (BC 2.7.1.112) (SCFR)
DE (Proto-oncogene tyrosine-protein kinase Kit) (c-kit).
GN KIT.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
(1)
SEQUENCE FROM N.A.
MEDLINE=99142897; PubMed=9989791;
Ma Y., Longley B.J., Wang X., Blount J.L., Langley K., Caughey G.H.;
"Clustering of activating mutations in c-KIT's juxtamembrane coding
region in canine mast cell neoplasms.";
J. Invest. Dermatol. 112:165-170(1999).
CC -!- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
CC 3-KINASE (PI3K).
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE CSP-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/
or send an email to license@isb-sib.ch).
-----
EMBL; AF044249; AAD02327.1; -.
DR HSP; P11362; LFCK
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001824; RTKinaseIII.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00047; Ig; 2.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 2.
DR SMART; SM00409; IG; 3.

```

```

SMART; SM00219; TyKc; 1.
DR PROSITE; PS00835; IG_LIKE; 2.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00109; PROTEIN KINASE TVR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
KW Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
KW Immunoglobulin domain.
FT SIGNAL 1 24
FT CHAIN 25 975
FT DOMAIN 25 519
FT TRANSMEM 520 542
FT DOMAIN 543 975
FT DOMAIN 588 936
FT NP_BIND 594 602
FT BINDING 622 622
FT ACT_SITE 791 791
FT MOD_RES 822 822
FT CARBOHYD 96 96
FT CARBOHYD 132 132
FT CARBOHYD 147 147
FT CARBOHYD 286 286
FT CARBOHYD 296 296
FT CARBOHYD 303 303
FT CARBOHYD 355 355
FT CARBOHYD 370 370
FT CARBOHYD 403 403
FT CARBOHYD 466 466
FT CARBOHYD 489 489
SQ SEQUENCE 975 AA; 109335 MW; 8F570BDB9F05B1CB CRC64;
Query Match 22.9%; Score 1207; DB 1; Length 975;
Best Local Similarity 31.9%; Pred. No. 2.5e-73;
Matches 330; Conservative 165; Mismatches 375; Indels 166; Gaps 34;
Qy 31 LPVTKVLISHENNGSSAGPSSVYRVGSPEDLQCTPRQSEGVTVAAATVEVAESGI 90
Db 11 LCVLLLLLLLVGRTGSS--QPS-----VSFGPELSPSIHPAKSELIVSGVD- 54
Qy 91 TLQVQLATPGDLSCLVWPKHSSLGCPHFQDLQNRGIVSMALNTVTQAGEYLLHIQSER 150
Db 55 ELRUSCTDTPGVK--WTFE--TLG-QUNENHNEWITEKA-----EAGTGYTCTNR 102
Qy 151 ANYTVLETVNVRD--TQLYVLRPPYFRKMNODALLCISGVPEPTVEVWVLCSSHRESCKE 209
Db 103 DGLRSIYVVRDPKALFLVDLPYKGEKNDTLVRC-----PLTDPEVINYSLRGC- 153
Qy 210 EGPVAVRKEEKVLHELFGTDIR-----C--CARNALGRE-CTKLFTIDLNOAPQS- 256
Db 154 EGKPLKDLTFVADPKAGITIRNVKREYHRLCHCSADQKGTVLTKKFTLVKRAAIRAV 213
Qy 257 -----TLPLQFLKVGPEPLWTRCKAIHVNHGFLTWELDKALEBGSYFEMSTYNTMTI 311
Db 214 PWSVSKTSLLEKEGAFVSMCFIKDVSSFFVDSMMWIKENSQOTNAQTSNWHHGDFFNE 273
Qy 312 RILLAFVSSVGRNDTYGTCSSSKHPSQSALVTILE---KGFINA-TSSQBEYEIDPYEK 367
Db 274 ROEKLITISSARVNSGVFMCVANTTILEVVDKGFINIFPMWMTITFVNDQGN 333
Qy 368 FCFSVRFKAYPR-IRCTWIFSQASF--PCRGLEDGYSISKFDH-----KNKPEY 417
Db 334 VDLIVEYEAYPEHEHQQWYVMNRTFTDKWEDYPKSDNESIRYSELHLTRLKNGEGTY 393
Qy 418 IFYAENDDAQTKMFTLNIRKKPOVLANASQA--SCSDGYPLPSTWTKCKSDKSPNC 475
Db 394 TFQVNSDVNSSVTFNVVYNTKPEILTHESUTNGMLQCQVAGFPFPAVGVFCPOABORC 453
Qy 476 TEEL-PEGVWKNKARVKFGQWSSSTLNMSEAGKGLLVKCCAVNSMGTCSETIFLNSPG 534
Db 454 SVPIGPMQVQNSLSLSPGKLVQSSIDYSAFKHNGTVECRAYNVNVRG--SAFEN--- 508
Qy 535 PFPFIQDNISFYATIGLCLPFIIV-----VLIVLICHYKKQFRYESOLOMI-QVTGPLD 587

```



Db 509 -FAFEQIHPHTLFTLLIGFVIAAGMCMIIIMILTYKVLQKPMYEVQKVVVEING--- 564  
 QY 588 NEYFVDFROVEXDLKWFPPRENLEFGKVLGSGAFGRVWNAATAYGISTGVSIVAVKML 647  
 Db 565 NNVYIIDPTQLPYDHKWEFPNRLSFGKTLGAGAFGKVEATAYGLIKSDAAMTVAVKML 624  
 QY 648 KERADSCKEALMSLKMTHLGHNDINVLGACTLGGPVLYLFEYCCYGDLLNLYLRSK 707  
 Db 625 KPSAHLTEREALMSLKVSLVGLNHMIVNLGACTVGGPTLVITYECCYGDLLNLYLRSK 684  
 QY 708 REXPHRTWTEIFKEHNFSSYPTFOAHNSMSPGSEVOLHPLDQLSGFNGNSIHSEDEI 767  
 Db 685 RDSFICKQK---DHG-----EVALYK-----NLLHSKSS 712  
 QY 768 -----EYENQK-----RLAEEDLDLN-----VLTFFDLICFA 795  
 Db 713 CSSTNEYMDMKPGVSYVPTKADKRSARITGSYIERDVTPIAMEDDELALDLELLSFS 772  
 QY 796 YQVAKGMFELEFKSVHRDLAARNVLVTHGKVKICDFGLARDILSDSSVYVRCNARLPV 855  
 Db 773 YQVAKGMFLASKNCIHRDLAARNILVTHGKVKICDFGLARDIKDNDSSVYVRCNARLPV 832  
 QY 856 KWMAPESLFEGIYTIKSDVMSYGILLWEIFSLGVNPPGIPVDANFYKLIOSGFKMQPFF 915  
 Db 833 KWMAPESIFNCVYTFESDVMYSYGIFLWELFSLGSPYFGMPVDSKFMVMEKGFRLMSPE 892  
 QY 916 YATEGIYVQMSWAFDSRKPSPNITSLFGCOLAEAEACIRTSIHLPKQAAPQORGG 975  
 Db 893 HAPAEYDYMKTCDWADPLKRPFTKQIVLQIEKQISDSTN-----HIYNSLAN----- 940  
 QY 976 LRAQSPORQVYKIHRR 991  
 Db 941 -CSPNPERPVVDHSVR 955

## RESULT 12

KIT\_HUMAN  
 ID KIT\_HUMAN STANDARD; PRT; 976 AA.  
 AC Fl0721;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-JUL-1989 (Rel. 11, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)  
 DE (Proto-oncogene tyrosine-protein kinase Kit) (c-kit) (CD117 antigen).  
 GN KIT.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Fetal brain, and Term placenta;  
 RX MEDLINE=88111521; PubMed=2448137;  
 RA Yarden Y., Kuang W.-J., Yang-Feng T., Coussens L., Mumentau S.,  
 RA Dull T.J., Chen E., Schlessinger J., Francke U., Ullrich A.;  
 RT "Human proto-oncogene c-kit: a new cell surface receptor tyrosine  
 RT kinase for an unidentified ligand.";  
 RL EMBO J. 6:3341-3351(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93064697; PubMed=1279499;  
 RA Giebel L.B., Strunk K.M., Holmes S.A., Spritz R.A.;  
 RT "Organization and nucleotide sequence of the human KIT (mast/stem  
 RT cell growth factor receptor) proto-oncogene.";  
 RL Oncogene 7:2207-2217(1992).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97179223; PubMed=9027509;  
 RA Andre C., Hampe A., Lachaume P., Martin E., Wang X.P., Manus V.,  
 RA Hu W.X., Galibert F.;  
 RT "Sequence analysis of two genomic regions containing the KIT and the  
 RT PMS receptor tyrosine kinase genes.";  
 RN [4]  
 RL Genomics 39:216-226(1997).  
 RP VARIANT LYS-583.  
 RX MEDLINE=92291284; PubMed=1376329;  
 RA Fleischman R.A.;  
 RT "Human piebald trait resulting from a dominant negative mutant allele  
 RT of the c-kit membrane receptor gene.";  
 RL J. Clin. Invest. 89:1713-1717(1992).  
 RN [5]  
 RP VARIANT LEU-584.  
 RX MEDLINE=92133600; PubMed=1370874;  
 RA Spritz R.A., Giebel L.B., Holmes S.A.;  
 RT "Dominant negative and loss of function mutations of the c-kit  
 RT (mast/stem cell growth factor receptor) proto-oncogene in human  
 RT piebaldism.";  
 RL Am. J. Hum. Genet. 50:261-269(1992).  
 RN [6]  
 RP VARIANT ARG-664.  
 RX MEDLINE=92020918; PubMed=1717985;  
 RA Giebel L.B., Spritz R.A.;  
 RT "Mutation of the KIT (mast/stem cell growth factor receptor)  
 RT protooncogene in human piebaldism.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:8696-8699(1991).  
 RN [7]  
 RP VARIANT VAL-816.  
 RX MEDLINE=94013473; PubMed=7691885;  
 RA Furitsu T., Tsujimura T., Tono T., Ikeda H., Kitayama H.,  
 RA Koshimizu U., Sugahara H., Butterfield J.H., Ashman L.K.,  
 RA Kanayama Y., Matsuzawa Y., Kitamura Y., Kanakura Y.;  
 RT "Identification of mutations in the coding sequence of the proto-  
 RT oncogene c-kit in a human mast cell leukemia cell line causing  
 RT ligand-independent activation of c-kit product.";  
 RL J. Clin. Invest. 92:1736-1744(1993).  
 RN [8]  
 RP VARIANTS PIEBALDISM GLY-791 AND VAL-812.  
 RX MEDLINE=93322624; PubMed=7687267;  
 RA Spritz R.A., Holmes S.A., Itin P., Kuester W.;  
 RT "Novel mutations of the KIT (mast/stem cell growth factor receptor)  
 RT proto-oncogene in human piebaldism.";  
 RL J. Invest. Dermatol. 101:22-25(1993).  
 RN [9]  
 RP VARIANT PIEBALDISM 893-GLU--PRO-896 DEL.  
 RX MEDLINE=96287384; PubMed=8680409;  
 RA Riva P., Milani N., Gandolfi P., Larizza L.;  
 RT "A 12-bp deletion (7818del12) in the c-kit protooncogene in a large  
 RT Italian kindred with piebaldism.";  
 RL Hum. Mutat. 6:343-345(1995).  
 RN [10]  
 RP VARIANT GIST VAL-559 DEL.  
 RX MEDLINE=98361155; PubMed=9697690;  
 RA Nishida T., Hirota S., Taniguchi M., Hashimoto K., Isozaki K.,  
 RA Nakamura H., Kanakura Y., Tanaka T., Takabayashi A., Matsuda H.,  
 RA Kitamura Y.;  
 RT "Familial gastrointestinal stromal tumours with germline mutation of  
 RT the KIT gene.";  
 RL Nat. Genet. 19:323-324(1998).  
 CC -1- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL  
 CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.  
 CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND  
 CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL  
 CC 3-KINASE (PI3K).  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- DISEASE: DEFECTS IN KIT ARE THE CAUSE OF A DEVELOPMENTAL  
 CC ABNORMALITY KNOWN AS PIEBALDISM. AN AUTOSOMAL DOMINANT GENETIC  
 CC DISORDER OF PIGMENTATION CHARACTERIZED BY CONGENITAL PATCHES OF  
 CC WHITE SKIN AND HAIR THAT LACK MELANOCYTES.  
 CC -1- DISEASE: DEFECTS IN KIT ARE THE CAUSE OF A GASTROINTESTINAL  
 CC STROMAL TUMOR (GIST).  
 CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
 CC PROTEIN KINASES.  
 CC -1- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.



662 IGGTFLVITECCYGGDLNLFRRKDRSPICSKQEDHAEALYKNLLHSSKSSCDSTNEY 721  
744 VQLHPPLDQL---SGFNCSHSEDEIYENQKLAEEEDNVLTPEDLLCFAYQVAK 800  
722 MDMPKGVSYVPTKADKERSVRISYIERDVTVPALMEDEALDL---EDLLSFSYQVAK 778  
801 GMEFLEPKSCVHRDLAARNLVTHGKVKICDGLARDILSDSSVYVRGNARLPVKWAP 860  
779 GMAFLASKNCIHRDLAARNILLTHGRITKICDGLARDIKDKNVYVKGARLPVKWAP 838  
861 ESLEFEGYITKSDVMSYGILLWELFSLGVNPGIPVDANPYKLIQSGFQWQEPFYATEG 920  
839 ESIFNCVYTFESDVMWSYGIWELFSLGSSPYGMPVDKSFYKMKEGFRMLSPHAPAE 898  
921 IYFVMSQWAFDSCRKSPFNLTSLFLGCOLAE 953  
899 MYDIMKTCWADPLKRFTRFQIVOLIEKQISES 931

RESULT 13  
KIT\_CHICK  
ID KIT\_CHICK STANDARD; PRT: 960 AA.  
AC Q08156;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)  
DE (Proto-oncogene tyrosine-protein kinase Kit) (c-kit).  
GN KIT.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=White leghorn; TISSUE=Brain;  
RX MEDLINE=9329295; PubMed=7685729;  
RA Sasaki E., Okamura H., Chikamune T., Kanai Y., Watanabe M.,  
RA Naito M., Sakurai M.;  
RT "Cloning and expression of the chicken c-kit proto-oncogene";  
RL Gene 128:257-261(1993).  
CC -!- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL  
CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.  
CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND  
CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL  
CC 3-KINASE (PI3K).  
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
CC tyrosine phosphate.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: HIGH IN THE BRAIN AND TESTES AND ALSO  
CC PRESENT IN THE BURSA OF FABRICUS, HEART, KIDNEY, LUNG, SPLEEN  
CC THYMUS AND OVARY.  
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
CC PROTEIN KINASES.  
CC -!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))

EMBL: D13225; BAA02506.1; -;  
PIR: JN0677; JN0677.  
HSSP: P11362; 1FGK.  
InterPro: IPR007110; Ig-like.  
InterPro: IPR003598; Ig\_c2.  
InterPro: IPR00719; Prot kinase.  
InterPro: IPR001824; RTKinaseIII.

InterPro: IPR001245; Tyr\_kinase.  
Pfam: PF00047; Ig; 3.  
Pfam: PF00069; pkinase; 1.  
ProDom: PD000001; Prot kinase; 2.  
SMART: SM00408; IGC2; 1.  
SMART: SM00219; TyrKc; 1.  
PROSITE: PS50835; IG LIKE; 1.  
PROSITE: PS00107; PROTEIN KINASE ATP; 1.  
PROSITE: PS00111; PROTEIN KINASE DOM; 1.  
PROSITE: PS00109; PROTEIN KINASE TYR; 1.  
PROSITE: PS00240; RECEPTOR TYR KIN III; 1.  
Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;  
Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;  
Immunoglobulin domain.  
FT SIGNAL 1 24  
FT CHAIN 25 960  
FT DOMAIN 25 505  
FT TRANSMEM 506 530  
FT DOMAIN 531 960  
FT DOMAIN 575 913  
FT NP BIND 581 589  
FT BINDING 609 609  
FT ACT SITE 777 777  
FT MOD RES 808 808  
FT CARBOHYD 76 76  
FT CARBOHYD 135 135  
FT CARBOHYD 149 149  
FT CARBOHYD 269 269  
FT CARBOHYD 286 286  
FT CARBOHYD 306 306  
FT CARBOHYD 318 318  
FT CARBOHYD 338 338  
FT CARBOHYD 343 343  
FT CARBOHYD 356 356  
FT CARBOHYD 453 453  
FT CARBOHYD 469 469  
SQ SEQUENCE 960 AA; 107311 MW; 0E93850527AB68F6 CRC64;

Query Match 22.4%; Score 1181; DB 1; Length 960;  
Best Local Similarity 31.8%; Pred. No. 1.4e-71;  
Matches 310; Conservative 159; Mismatches 356; Indels 150; Gaps 30;

QY 45 GSGAGKPSYRMVGRGSPEDLQCTPRQSEGTVEATVEAESGITLQVLATPGDLSL 104  
DB 24 GSVPHSESLVWKGELRLKCN-----EAGPTWTFQNSDPSAKTR 65  
QY 105 L-----WPKHSSLCQPHFDLQNRGIVSMALNVTETQAGEYLLHIQSERANYTLFTV 159  
DB 66 ISNEKEWHTKNAITIRDIGRECKSKG-----SIVN-----SFYV 99  
QY 160 NVYRDTQLYVLRPRYFRKMNQDALLCISBGVPEPTVEWVLCSSHRESCKEKGPAVRKEE 219  
DB 100 FVKDENVLFLVDLSLYGKEDSDILLVCLPTDPP-VLNFTRKCDGKPLPKNMTFIPNPK 158  
QY 220 ----KVLHLEFGTDIRCCAR-NALGRECTKLPIDINQAP-QSTLPOL-----FLKVG 266  
DB 159 GIIIKNVQSRSPKGVQCLAKHNGVEKISEHIP--LNVRPVHKAFLPVITLSKYEELKEG 215  
QY 267 EPLWTRCAIHVNHGFGTWELEDKALBEGSYFEMSTYSTNRTMIRILLAFVSSVGRNDT 326  
DB 216 BEPEFTCIITDVSQVSKASWISYKSAIVTSKRNLDGYERK----LTLNIRSVGVNDS 271  
QY 327 GYITCSSSKHP--SOSALVTI--LEKGFINATSSQE-EYEDPYKFCFSVFAYPRIR 381  
DB 272 GEFTC-QHENPFGKTNATVTLKALAKGFVRFLATWNTTIDINAGQGNLTVEYAYPK-- 328  
QY 382 CTWIFSQASPFCEQGLEGDYGSISKFDHKNK-----PGEYIF 419  
DB 329 -----PKEEVMVMYMETLQNSDHYVYKFTVGNNSYSELHLTRLKGTGEGIYTF 378  
QY 420 YAENDDAQFTQFTLNTIRKKPOVLANASQ--ASCSSDGYPLPSWTWKCKSDKSPNCTE 477  
DB 379 FVNSDASSSVTFNVYVTKPEILTLDMGLNDILQCVATGFPAPTIYVYFCPTQRCLD 438





FT CONFLICT 519 519 S -> T (IN REF. 2).  
 SQ SEQUENCE 1088 AA; 122641 MW; 590C9BB0418801E7 CRC64;  
 Query Match 22.0%; Score 1157.5; DB 1; Length 1088;  
 Best Local Similarity 31.9%; Pred. No. 6.2e-70;  
 Matches 308; Conservative 162; Mismatches 328; Indels 169; Gaps 32;

QY	94	VOLATPGLDLSCLWVFKHSSLGCG-----PHFDL---QNRGIVSMALNVTETOAGE- 141
DB	38	VPLSSFSLRG---FGSEVSWQHPMSEBDDPNVEIRTEENNSSLFVTVLEVVNAAHT 94
QY	142	-----YLLHHQSRANYTVLFTVNRDQYVLRPP-----YFRKMNQDALLCISEG--- 189
DB	95	GWTCYNHTQTBESE-----IEGRHIYIV---PDPMAFVPLGMDTSLVIVEEDDSA 145
QY	190	-VP-----EPVEVWVLCSSHR-----ESKEEGP---AVRKEEKVLHELFGT 228
DB	146	IIICLTDPDTEVLHNGRLVPASVDSRGFNGTFSGVGYCEATVR-----GR 195
QY	229	DIRCCARNALGRCTKLTIDLNAQPOSTLPQLFLVKGPELWIRCKAIHNVHGFGLTWEL 288
DB	196	TFKTSFENVYALKATSELNLEMDTR-----QTVYKAGETIVVTC-AVFNEVVDLQWTY 248
QY	289	EDKALREGSYFEMSTYSTNRTMR-----ILLAFVSSVGR-----NDTGYTCS---SSKH 336
DB	249	PGEVRNKGI-----TWLEBKLPSEIKLVYTLTVPKATVKDSGDYECARQATKE 297
QY	337	PSOSALVTIL-BKGFINATSSQEEYE-IDPEYKFCFSVRFKAYPRICRTWIFSOASF-- 391
DB	298	VKEMKVTVTSVHEKGFVQIRPPTFGHLETVNLHQVREFVVEVQAYPTPRISWLKDNLTLE 357
QY	392	-----PCEQRLEDGY-SISKFCDHKNK-PGBYIFYAENDDAQFTMFTLNIRKPKQVL 443
DB	358	NLTEITTDVQSRQETRYOSKULKIRAKEEDSGHYTIVQVNDMDKSYTFELSLVPSIL 417
QY	444	-----ANASQASCSDDGYLPSWTW-----KKCSDK-----SPNCTEEIPEGVW 484
DB	418	ELVDHHGGGGGTQVRCETAGTFLPNIEMWMIKDKKCNNDTSWTVLASNVSNIIITE--F 475
QY	485	NKANRKFQGVQWSSSTLNNSEAGKLLVKCCAYNSMGTSCTFIENSGPPFPFIQDNIS 544
DB	476	HQGRSTVEGR-----VSFAKVEETIAVRCLAKNDLIGNRELKLVAFS-----LRSEIT 525
QY	545	FYATIGLCLPFIIVLIVLICHYKQFPRYESQLQMTQVTGPLDNEYFYVDFRDYEYDLKW 604
DB	526	VAAAVLVLLVIVSLIVLVVWKQPRYEIRWRVIESISPDGHEIYIYDPMQLPYDSRW 585
QY	605	EPRENLEFGKVLGSGAFGRVMNATAYIGSKTVGSQVAVKMLKEKADSCKEALMSELK 664
DB	586	EFPRDGLVGRILGSGAFGRVVEGTAYGLSRSPVMKVAVMKLPKPTARSSEKQALMSELK 645
QY	665	MMTHLGHHDNIIVNLGACTLGGPVYLFIFFCYCGDILLNLYRSKREF-----711
DB	646	IMTHLGHPLNIVNLGACTSGGFIYIITEYCYFGDLVNLHKRDSFMSRHPKPKKDLDD 705
QY	712	-----HRTWTEIFKEHNFSSYPTFOAHNSNSMP--GSREVOLHPPLDOLSGFNGN 759
DB	706	IFGLNPADESTRSVILSPENNGDYVDMKQADTTQVPEMLERKEVSKYSIDIQESLYDRPA 765
QY	760	SIHSEDEIEYENQKRLAEBEEDLNLVTFEDLLCFAYQYAKGMFELEPKSCVHRDLAARN 819
DB	766	SYKKKSMLDSEAKNLLSDDDEGLTLI---DLLSFTYQVARGMEFLASKNCVHRDLAARN 822
QY	820	VLVTHGVKVICDFGLARDITLSDSSVYVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGI 879
DB	823	VLLAQGKIVKICDFGLARDIMHDSVYVYSGSTFLPVKWWAPESIFDNLYTTLSDVMSYGV 882
QY	880	LLWEISLGVNYPGIPVDANFYKLIQSGFKMEQPFYATEGIFYVVMQSCWAFDSKRKPSF 939
DB	883	LLWEISLGTPTFGWAVDSTFYNKIKSGYRMAKPDHATSEVYIEMVQWNSEPEKPSF 942
QY	940	PNLTSFL 946



GenCore version 5.1.6  
Copyright (C) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2003, 07:20:28 ; Search time 83.4579 Seconds  
(without alignments)  
3067.272 Million cell updates/sec

Title: US-09-919-408a-2

Perfect score: 5264

Sequence: 1 MRALAQRSDRRLLLVLSV.....RGGLRAQSPQROVKIHRERS 992

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rident:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2099	39.9	406	11	Q8BR20
2	1268	24.1	976	13	Q8JFE5
3	1268	24.1	976	13	Q9W755
4	1239	23.5	984	13	Q8AXC6
5	1234.5	23.5	974	11	Q63702
6	1232.5	23.4	978	11	Q63116
7	1225	23.3	979	11	Q8C8K9
8	1220	23.2	954	13	Q91909
9	1215.5	23.1	974	13	Q98S03
10	1215.5	23.1	977	13	Q98S01
11	1214	23.1	977	13	Q918N6
12	1209.5	23.0	979	6	Q8WN23
13	1208	22.9	964	6	Q97744
14	1207	22.9	964	6	Q9TQ01
15	1207	22.9	964	6	Q9TQ00
16	1205	22.9	978	6	Q9XS93

17	1204.5	22.9	977	13	Q98SU2	Q98su2 danio kerri
18	1201	22.8	948	6	Q9TID7	Q9ttd7 trichosurus
19	1198.5	22.8	977	13	Q98S04	Q98su4 danio albol
20	1197.5	22.7	992	13	Q8AXU0	Q8axu0 oncorhynch
21	1194	22.7	972	4	Q99662	Q99662 homo sapien
22	1188.5	22.6	975	13	P79750	P79750 fugu rubrip
23	1173	22.3	945	6	O77589	O77589 equus cabal
24	1170.5	22.2	724	6	Q9MYN0	Q9myn0 bos taurus
25	1139.5	21.6	1062	13	Q8AXC7	Q8axc7 fugu rubrip
26	1136	21.6	1059	13	Q9DE49	Q9de49 brachydanio
27	1135.5	21.6	1078	13	Q8AXC8	Q8axc8 fugu rubrip
28	1126.5	21.4	1087	13	Q9PUF6	Q9puf6 gallus gall
29	1110	21.1	1097	11	Q8R406	Q8r406 rattus norv
30	1079	20.5	1106	4	Q8N5L4	Q8n5l4 homo sapien
31	1073	20.4	1019	13	Q8UVR8	Q8uvr8 fugu rubrip
32	1060	20.1	1048	13	P79749	P79749 fugu rubrip
33	1013.5	19.3	986	13	Q8UVR9	Q8uvr9 fugu rubrip
34	997.5	18.9	923	6	O97745	O97745 sus scrofa
35	980	18.6	1379	13	P79701	P79701 coturnix co
36	974	18.5	1327	13	Q9QHL3	Q9qhl3 gallus gall
37	961	18.3	1363	11	Q91ZT1	Q91zt1 rattus norv
38	957	18.2	1301	13	Q8UWU9	Q8uwu9 brachydanio
39	952.5	18.1	1302	13	Q8AXB3	Q8axb3 brachydanio
40	946	18.0	1173	13	Q9PTL0	Q9ptl0 brachydanio
41	939	17.8	563	11	Q925F7	Q925f7 rattus norv
42	937.5	17.8	1345	11	Q8VCD0	Q8vcd0 mus musculu
43	901.5	17.1	323	11	Q9EQ22	Q9eq22 rattus norv
44	877.5	16.7	323	11	Q9EQ24	Q9eq24 rattus norv
45	833	15.8	345	13	Q9PVU7	Q9pvu7 lethenteron

#### ALIGNMENTS

#### RESULT 1

Q8BR20 PRELIMINARY; PRT; 406 AA.

AC Q8BR20;  
DT 01-MAR-2003 (Tremblrel. 23, Created)  
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
DE FMS-like tyrosine kinase 3.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Brain;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium,  
RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs."  
RL Nature 420:563-573(2002).  
DR EMBL; AK045865; BAC32514.1; ..  
SQ SEQUENCE 406 AA; 45834 MW; 155394D1676D7D37 CRC64;

Query Match 39.9%; Score 2099; DB 11; Length 406;  
Best Local Similarity 99.0%; Pred. No. 9e-176;  
Matches 401; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy	1	MRALAQRSDRRLLLVLSVMILEVTNODLPVTKCVLISHENNGSSAGKPSRYMVRGS	60
Db	1	MRALAQRSDRRLLLVLSVMILEVTNQDLPVTKCVLISHENNGSSAGKPSRYMVRGS	60
Qy	61	PEDLQCTPRRSEGTYEATVAESGSIITLQVLATPGDLSPGLVFKHSLGQCPHFD	120
Db	61	PEDLQCAPRRSEGTYEATVAESGSIITLQVLATPGDLSPGLVFKHSLGQCPHFD	120
Qy	121	LQNRGIVSMALINVTETQAGEYLLHIQSERANYTLFTVNVVVDLTQYLVRFPFRMENO	180
Db	121	LQNRGIVSMALINVTETQAGEYLLHIQSEAAANYTLFTVNVVVDLTQYLVRFPFRMENO	180





```

QY 386 FQASFPCEQRLGLEDYSISKFDH-----KNKPEGYIFYAEND 424
Db 344 FM-----GRGLHNTSDHVTTRSHSVYTSSELKVLRLKMSGGVYTFQASNG 390
QY 425 DAQFTKMTFLNTRKKQVLANASA--SQACSSDGYPLPSWTWKCKSDKSPNCTEEL--- 479
Db 391 DASVNTFTFIVISKPEIIVSHGEPVDGQVRCVAEGFPAPQITWYCEQPVARCQOQVAT 450
QY 480 --PEGVNNKANRKFVGQWSSSTLNNSEAGKLLVKKCCAYNSMGTSCETIFLNSPGPF- 536
Db 451 QEQNVITVTLSPFLGKTEVESRVNIS-GRSTILECVA-TVEGQAFTLFSISERTIS 508
QY 537 -----PFIQNTISFYATIGLCLPFIVVLIVLICHYKKQFRYESQLQMTQVGPLONEVF 591
Db 509 HDLFSPLLIGSVS--AACILCL-----ILVLF-YKTMQPKYQIQWKVIE--GIHGNVY 559
QY 592 YVDFRDEYDLKWEFPRENLEFKVLGSGAGFVGMNATAYIGSKTGVSTQVAVKMLKEKA 651
Db 560 YIDPTQLPYDHOHEFPKRNLFPGKTLGSGAFGKVEATAYAGLANEDSMMTVAVKMLKSA 619
QY 652 DSCKEALMSELKMMTHLGHHDNIVNLGACTLSGPVYLIFEYCCYGDLLNLYRSKREKF 711
Db 620 HSTKEKALMSELKVLIVLGNHINIVNLGNCTVGGPTLVITEYCCFGDLLNFLRRRESF 679
QY 712 HRTWTEIFKEHNFSSYPTFOAHNSMSPGSRREVQLHPL--DQLSGF-----NGN 759
Db 680 -----ICFKLEEDCHY-----RNIMLQREMGDSLNGVMTWRPSAAGKPS 720
QY 760 STHSEDEIYENOKRLABEE-----EEDLNVTFDLLCFAYOVAKGMEFLEFKCVHR 813
Db 721 SSSSEKRRSLREGSPVEEDESSEMFDEDSLSDTEDLSFSYQVAKGMEFLEFKCNCHR 780
QY 814 DIAARNVLVTHGVKVKICDFGLARDILSDSSYVVRGNARLPVKWMAPESLFEGIYTIKSD 873
Db 781 DIAARNILTOGRVAKICDFGLARDINTDSNYVVRGNARLPVKWMAPESLFEGIYTFESD 840
QY 874 VMSYGILLWEIFSLGNVPYGPVVDANFYKLIOSGFKMEQFFYATIGIYFVMSQWAFDS 933
Db 841 VMSYGILLWEIFSLGNSPYGPMVDKAFYKLIKEGYRMDAPEFAPSEMYQIMESCDADP 900
QY 934 RKRPSFNLTSFLGCOLAE-ABEACTRISHLPKQAAPQORGLRAQS----- 980
Db 901 LNRPPKRVKVERIEQQLSDTTKHIYLFNSRVPMRGRRESSTHSMASQFFNSAGNNSP 960
QY 981 PORQVKIHE 990
Db 961 PSRLPLLHE 970

```

## RESULT 5

```

Q63702
ID Q63702 PRELIMINARY; PRT; 974 AA.
AC Q63702;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE C-kit receptor tyrosine kinase isoform.
OS Rattus rattus (Black rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BN/fmai.
RA Tsujimura T., Tono T., Yamazaki M., Nomura S., Kitamura Y.;
RT "Two isoforms of rat c-kit receptor tyrosine kinase.";
RL Nucleic Acids Res. 0:0-0(0).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
DR EMBL; X62491; CAA44354.1; -.
DR HSSP; P11362; 1FGK.

```

```

DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001824; RTKinaseIII.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 2.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS00835; IG LIKE; 3.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW Phosphorylation; Receptor; Transferase; Transmembrane;
KW Tyrosine-protein kinase.
SQ SEQUENCE 974 AA; 108955 MW; B9BFAD9BA0DA190D CRC64;

```

Query Match 23.5%; Score 1234.5; DB 11; Length 974;  
 Best Local Similarity 32.3%; Pred. No. 3.9e-99;  
 Matches 323; Conservative 176; Mismatches 340; Indels 161; Gaps 37;

```

QY 36 CVLISHENNGSAGKPSYRMVGRGSPEDLOCTPRROSEGTIVYEAATVEVAESGITLQVQ 95
Db 12 CVLLVLLRGQTGTGSPSA-----SPGEPSPPSIQPAQSELIVEAGD-TIRLT 57
QY 96 LATPGDGLCLWFKHSSLCQPHFDLQNRGIVSMALNV-TETQAGEYLLHISERANYT 154
Db 58 CTDPAFVK--WTFE-----ILDVRIENKQSEWIR--EKAETAT 92
QY 155 VLFT-----VNRD-TOLVLRPRFKMENQDALLCISEGVPEPTV-EWVLC 200
Db 93 GKTCVSGSLRSSIVFVRDPAVLFLVGLPLFGKEDNALVRC---PLTDPQVSNYSLI 149
QY 201 SSHRESC-----KEGPVAVRKEEKLHELFGTDIRCCARNALGRECTKLTIDLNQ 252
Db 150 ECDGKSLPTDLKFVPNPKAGITIKNVKRAVHRLC--IRCAAQREGKWMRSDFTLKVRA 206
QY 253 A-----POSTLPQL--FLKVGPELWIRCKAIHVNHGFLTW-----ELEDKALEE 295
Db 207 AIKAIPIVSVSPETSHLLKEGDTFTVICTIKDSTVSDSMWIKLNPOFQSKAQKRNWSHQ 266
QY 296 GSYFEMSTYSTNRMLRILLAFVSSVGRNDTGYTCTSSSKHPSQSALVT---ILEKGFN 352
Db 267 GDF-----NYERQETLT-----ISSARVNDSGVMFYANNFTGSANVTTLTKVVEKGFN 316
QY 353 --ATSSQBEYEIDPYEKFCFSVRFKAYPR-IRCTWIFSQASFPCEQRLG-----DGYIS 405
Db 317 IIPVKNTTVFVTDG-ENVDLVVEFEAYPKHQQWIMYMRNT--PTNRGEDYVKSDNQSI 373
QY 406 KPCDH-----KNKPEGYIFYAENDDAOFTKMTFLNIRKKQVLA--NASASQASCS 456
Db 374 RYVNEELRLRLKGTGGTYTFLVNSGDSVASVTFDVYVNTKPEILTYDRLMNGRLQCVAA 433
QY 457 GYPLPSMTWKCKSDKSPNCTEETIPE-GVNNKANRKFVGQWSSSTLNNSEAGKLLVKC 515
Db 434 GFPEPTIDWYFCGAEQRCRTVPVPPVDVQI ONASVSPFGKLVVQSSIDSSVFRHNGTVEC 493
QY 516 CAYNSMGTSCTETIFLNSPGPF--IQDN-----ISFYATIGLCLPFIVVLIVLICH 565
Db 494 KASNAVGS--SAFFN---FAPKEIQPHLTPLLIGFVVVTAGL---MGIIVMVLAY 543
QY 566 KYKQFRYESQLQMI-QVTGPLDNEYVDYDRDYDLKWEFPRENLEFKVIGSGAFGR 624
Db 544 KYLQKPMYEQWKVVEEING---NNVYIDPTQLPYDHLKWEFPRENRLSFGKTLGAGAFGK 600
QY 625 VMNATAYGSKTGVSTQVAVKMLKEKADSCKEALMSELKMMTHLGHHDNIVNLGACTL 684
Db 601 VVEATAYGLIKSDAAMTVAVRMLKPSAHLTERBALMSELKVLVSLGNHMINVLLGACTV 660

```

QY 685 SGVYLIFCYCCYGLNLYRSREKFKHRTWTETIFKEHNFSSYPTFOAHNSNMPGSGREV 744  
 DB 661 GGPTLVITEYCCYGLNLYRKRKDSF - IFSKQEEQADAAALYKNLLSHKSSCDSSNEY 718  
 QY 745 QLHPPLDQLSGFN-----GNSIHSDEIEYENQKRLAEBEEDLNVLTFEDLLCFA 795  
 DB 719 -----MDMKPGSVYVPTKDKRSARIDSYIERDVTTPAIMEDELALDL---EDLLSFS 770  
 QY 796 YQVAKGMEFELEFKSCVHRDLAARNVLTGKVKIKCDFGLARDITLSDSVYVVRGNARLPV 855  
 DB 771 YQVAKGMAFLASNCVHRDLAARNVLTGKVKIKCDFGLARDITLSDSVYVVRGNARLPV 830  
 QY 856 KMAPSLFEGYITIKSDVMSYGILLWEIFSLGVNYPGIPVDANFYKLIQSGFQMEQPF 915  
 DB 831 KMAPSIFENCVYTFESDVMSYGIFLWELFSLGSSPYGMPVDSKFKYKMKIEGFRMLSP 890  
 QY 916 YATEGYIFVMSQWAFDSRKRFPNLTSLFGLQALAEAE 955  
 DB 891 HAPAAVYEVNMTKCDADPLKRPFTFKQVQLIEKQISDSSK 930

RESULT 6

ID Q63116 PRELIMINARY; PRT; 978 AA.  
 AC Q63116,  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE C-kit receptor tyrosine kinase.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley;  
 RX MEDLINE=92003944; PubMed=1912577;  
 RA Tsujimura T., Hirota S., Nomura S., Niwa Y., Yamazaki M., Tono T.,  
 RA Morii E., Kim H., Kondo K., Nishimune Y., Kitamura Y.,  
 RT "Characterization of Ws mutant allele of rats: A 12-base deletion in  
 RT tyrosine kinase domain of c-kit gene."  
 RL Blood 78:1942-1946 (1991).  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
 CC PROTEIN KINASES.  
 DR EMBL; D12524; BAA02094.1; -.  
 DR HSP; P11362; IFGK.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig c2.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR001824; RTK\_kinase.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR Pfam; PF00047; Ig; 1.  
 DR Pfam; PF00069; pkinase; 1.  
 DR ProDom; PD000001; Prot\_kinase; 2.  
 DR SMART; SM00408; IGC2; 1.  
 DR SMART; SM00219; TyrKc; 1.  
 DR PROSITE; PS00835; IG LIKE; 3.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.  
 DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
 DR ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;  
 KW Phosphorylation; Receptor; Transmembrane;  
 KW Tyrosine-protein kinase.  
 SQ SEQUENCE 978 AA; 109341 MW; 0958C33F19889051 CRC64;

Query Match 23.4%; Score 1232.5; DB 11; Length 978;  
 Best Local Similarity 32.2%; Pred. No. 5.8e-99;  
 Matches 323; Conservative 176; Mismatches 340; Indels 165; Gaps 37;  
 QY 36 CVLISHENNSSACKPSYRVRGSPEDLQCTPRQSEGVYEAATVEVAESGSIITLQVQ 95

DB 12 CVLILVLRGQTGTSQPSA-----SPGEPPPSIQPAQSELIVLEAGD-TIRLT 57  
 QY 96 LATPGDLSCLVWFKHSSLCGQPHFDLQNRGIVSNAILNV-TETQAGEYLLHIQSERANYT 154  
 DB 58 CTDPAFVK--WTFE-----ILDVRIENKQSEWIR--EKAETH 92  
 QY 155 VLFT-----VVRD--TQLYVLRPYPFRMENQDALLCISGEVPEPTV-EWYLC 200  
 DB 93 GKTYCVSGSLRSIYVVRDPAVFLVGLFLFGKEDNDALVRC---PLTDPQVSNYSLI 149  
 QY 201 SSHRESC-----KEGPAVVRKEEVLHEFGTDIRCCARNALGRECTKLTIDLNQ 252  
 DB 150 ECDGKSLPTDLKFPVMPKAGITIKNVRAYRLC---IRCAAQREKGMRSKDFETLVKVA 206  
 QY 253 A-----PQSLPQL--FLKVGELWIRCKALVHNGFGLTW-----ELDDKALEE 295  
 DB 207 AIKAIPIVSVSPETSHLKEGDTFTVICTIKDVSTVDSMWIKLNPQQAQVKNRSHWQ 266  
 QY 296 GSYFEMSTYSTNRTMIRIILAFVSSVGRNDTYTSCSSKHPQSALVT---ILEKGFN 352  
 DB 267 GDP---NYERQETLT-----ISSARVNDSGVFMCVANNTPFGSANVTTLTKVVEKGFN 316  
 QY 353 --ATSSQEEYEIDPYEKFCSYRFAKAYPR-IRCTWIFSQASFPCEORGLE----DGYGIS 405  
 DB 317 IPFVKNTTVFTDG-ENVDLVVEFEAYPKPEHQOIMYMRN--PTNRGEDIYKSDNQSN 373  
 QY 406 KFCDH-----KNKPGEYIFVAENDDAOFTKMTFLNIRKQPOLA--NASASQASCSSD 456  
 DB 374 RYVNELRRLTRLKGTEGTYTFLVNSNDVASVTFDYVNTKPEILTYDRLMNGRLQCVAA 433  
 QY 457 GYPLPSWTWKKGSDKSPNCTEPIE-GVWNKANKRVFGQWSSSTLNSEAGKGLLVKC 515  
 DB 434 GFPEPTIDWYFCTGAEQRCCTVPVPPVDVQIQNASVSPFKLVVQSSIDSVFRHNGTVEC 493  
 QY 516 CAYNSMGTSCTIFLNSPGPPF-----IQDN-----ISFYATIGICLPFVVLIIV 561  
 DB 494 KASNAVGKS--SAFFN---FAFKGNSKEIQIPHTLFTPLLIGFVVVTAAGL---MGIIVM 543  
 QY 562 LICHVKYKQFRYESQLOMI-QVTGPLDNEYFYVDYDYLKWEFPRENLEFGKVLGSG 620  
 DB 544 VLAYKILQKPMYEVQWVVEING---NNYVYIDPTQLPYDHKWEPFRNLSFGKTLGAG 600  
 QY 621 AFRVYNATAYGISKTGVSIOVAVKMLKEKADSCKEALMSLKMTHLGHHDNIYNLLG 680  
 DB 601 AFGKVVEATAGLIKSDAAMTVAVKMLPSAHLTEREALMSLKVLSYLGNEHNIYNLLG 660  
 QY 681 ACTLSGPVYLIFEYCCYGLNLYRSREKFKHRTWTETIFKEHNFSSYPTFOAHNSNMPG 740  
 DB 661 ACTVGGPTLVITEYCCYGLNLYRKRKDSF--IFSQEEQADAAALYKNLLSHKSSCDSS 718  
 QY 741 SREVQLHPPLDQLSGFN-----GNSIHSDEIEYENQKRLAEBEEDLNVLTFEDL 791  
 DB 719 SNEY-----MDMKPGSVYVPTKDKRSARIDSYIERDVTTPAIMEDELALDL---EDL 770  
 QY 792 LCFAYQVAKGMEFELEFKSCVHRDLAARNVLTGKVKIKCDFGLARDITLSDSVYVVRGNA 851  
 DB 771 LSPSYQVAKGMAFLASNCVHRDLAARNVLTGKVKIKCDFGLARDITLSDSVYVVRGNA 830  
 QY 852 RLPVKWMAPESLFEGYITIKSDVMSYGILLWEIFSLGVNYPGIPVDANFYKLIQSGFQ 911  
 DB 831 RLPVKWMAPESLFEGYITIKSDVMSYGILLWEIFSLGVNYPGIPVDANFYKLIQSGFQ 890  
 QY 912 EQPYATGEGYIFVMSQWAFDSRKRFPNLTSLFGLQALAEAE 955  
 DB 891 LSPHAPAAVYEVNMTKCDADPLKRPFTFKQVQLIEKQISDSSK 934

RESULT 7

Q8C8K9 PRELIMINARY; PRT; 979 AA.  
 ID Q8C8K9  
 AC Q8C8K9;  
 DT 01-MAR-2003 (Tremblrel. 23, Created)







```

QY 425 DAQFTKMTNIRKQVLAN-ASASQASCSGYPPLPSWTWKKCDKSPNCTEE----- 478
Db 401 VKSASITFDIKMTKEVARVKWENVTTLSRSYGPAPSLMTYQCTGIRTTCPNTTDLQ 460
QY 479 -----IPEGVWKKANRKFQGVSSSTPLNMSKAGLLVKCCAVNSMGTSCEITFLN 531
Db 461 PIQTHGGVPEG-----ESFGVGVESVLTG-PNRRMTVVCAFNLVGQSGSDTFMSD 511
QY 532 SPGPPFFIQDNISFYATIGLCLPFIV--VLIVLICHKYKQPRYESQLOMQIVTGPLDNE 589
Db 512 -----VSDQIFTSACMGSTVAMVGLLLIFMIYKQKPRYERWKIIEATN--GNN 562
QY 590 YFVDFPRDEYDLKWFPPRENLEFQKVLGSGARVWNAATAYGISTGVSIQVAVKMK 649
Db 563 YTFIDPTQLPYNEKWEFPRDKLGLKTLGAGAFGVKVEATAYGLGKEDNITRVAVKMLKA 622
QY 650 KADSCKEALMSLKMTHLGHNDNINLLGACTLSPVVLIEFYCCYCGDLLNLYRSKRE 709
Db 623 SAHPDEREALMELKLSHLGQKNVNLGACTHGGPVLVITEYCHGDLNLFASKAE 682
QY 710 KPHRTWTEI--FKB-----HNFSSYPTFOAHNSNMPGS-----REVQLHPPLDQLS 754
Db 683 NFNLFVMTIPDPPEPMTYKNISTERMF-VRSDSGISSACSDHYLDMPVTSRPTLSSS 741
QY 755 GNGNSIHSEDEIYENQKRLAEBEEDNLVTFEDLLCFAYQVAKMGFELEFKSCVHRD 814
Db 742 -----ECQEDSWPLDMDLLRFSQVQAQLOFLAAKNCIHRD 778
QY 815 LAARNVLTHGVKVKICDFGLARDILSDSVYVVRGNARLPVKWMAPELPEGITYIKSDV 874
Db 779 VAARNVLLTNSVAKICDFGLARDIWNDSYVYVGNARLPVKWMAPELPEGITYIKSDV 838
QY 875 WSYGILLWEIFSLGVNVPICPVDANFYKLIQSGFRMEOPFYATEGIYFVMOQSCWAFDSR 934
Db 839 WSYGIMLWEIFSLGKSPYNILVDSFKYMKIKGYQMSRDPFAPPENYIMKMCWNLDA 898
QY 935 KPSPFNLTSLFCQLAEBE 955
Db 899 ERPTFSKISQIMQLGETPE 919

RESULT 10
Q98SU1 PRELIMINARY; PRT; 977 AA.
AC Q98SU1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Receptor tyrosine kinase Fms.
OS Danio nigrofasciatus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=144739;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21359118; PubMed=11466528;
RA Parichy D.M., Johnson S.L.;
RT "Zebrafish hybrids suggest genetic mechanisms for pigment pattern
diversification in Danio.";
RL Dev. Genes Evol. 211:319-328(2001).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
PROTEIN KINASES.
DR EMBL; AF324481; RAK15303.1; -.
DR HSP; F11362; IFGK.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig C2.
DR InterPro; IPR003006; Ig MHC.
DR InterPro; IPR000719; prot_kinase.
DR InterPro; IPR001824; RTKinaseII.
DR InterPro; IPR001245; Tyr_pkinase.
Pfam; PF00047; Ig; 4.

```

```

DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD00001; Prot_kinase; 2.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS0835; IG LIKE; 3.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
DR PROSITE; PS00240; RECEPTOR TYR KIN III; 1.
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW Phosphorylation; Receptor; Transferase; Transmembrane;
KW Tyrosine-protein kinase.
FT VARIANT 60 60 V -> L.
FT VARIANT 103 103 A -> S.
FT VARIANT 742 742 P -> S.
FT VARIANT 899 899 E -> D.
SQ SEQUENCE 977 AA; 109881 MW; CF54E129FAB10E3E CRC64;

Query Match 23.1%; Score 1215.5; DB 13; Length 977;
Best Local Similarity 33.5%; Pred. No. 1.8e-97;
Matches 334; Conservative 160; Mismatches 356; Indels 147; Gaps 35;

QY 54 YRMVRGSPEDLOCTPR-ROSEGVVYEAATVEVAESGITLQVCLATPGDLSCLW---VFK 109
Db 13 FQOVQWSE-----PIRLNSGAL--AGTDVILDGSG--PLQLVCEGDPVTFPPRAK 62
QY 110 HSLGQCPHFIDLQNRGIVSMALNVTETQAGEY-LLHIQSERANYTVLFTVNVVDTQ-LY 167
Db 63 H-----KRVISKEVGKIRSRVEKATVDTGTGYKCVYINGSDSNLASSVHVFRDVRVLF 117
QY 168 V-----LRPYPRKMNQDALLCISEGVPEPTVEVWVLCSSHRSCKEKGPAVVRKEKVL 222
Db 118 VSPSSSLR--YVRK-EGEDLLPCLLTDPDAT-DFTFRMDNGSAAPYGMNATDPKRGVL 173
QY 223 ----HELFTDTRCCARNALGRECTKLPDIDNQAQSTLPLFLK-----VGEPLWI 271
Db 174 IRNVHPGFNADYICSAIRIGAEKVKISFINVIQRLRFP-PYVYLKRNVEYVKLVGRLQI 232
QY 272 RCKAIHVNHGFLTWELEDKALEEGSYFEMSTVSTNRTMIRILLAFVSSVRGNDTGYTC 331
Db 233 SCTTNPNFYVNTWTHSSKRLPKPE--EKSTWEGDRLAIESILT-ISSVQLSDTGNITC 289
QY 332 SSKHPSQALVT---ILEKGFINAT---SSQBEY-----EIDPYKFCFSVRFKAYPRI 380
Db 290 TQNEAGANSSTQLLVDEPVIRLSPKLSKLTGRLSIEVSEGEDVDLGLVLEIAYPPL 349
QY 381 RC-TW---IFSOASPPCEOR--GLEGYSIKFCDHKN--KPGEYIFYAENDDAQFTKMF 432
Db 350 TSHTWETPTSHNASLP-ENRFYNHNDRYEALLFLKELNFEIEIGQYTLNVKNSMKASITF 408
QY 433 TLNIRKQVLAN-ASASQASCSGYPPLPSWTWKKCDKSPNCTEEI----PEGVWKK 487
Db 409 DIRMVTKPVARKWENVTTLSRSYGPAPSLMTYQCTGIRTTCPNTTDLQIQQTVE 468
QY 488 ANRKVFGVWSSSTLNMSEAGKLLVKKCCAVNSMGTSCEITFLNSPGPPFFIQDNISFYA 547
Db 469 FQKESFGAVGVESVLTG-PNRRMTVVCAFNLVGQSGDTSFMSD-----VSDQIFTS 520
QY 548 TIGLCLPFIV--VLIVLICHKYKQPRYESQLOMQIVTGPLDNEIYFVDFRDEYDLKWE 605
Db 521 MCGSTVAMVGLLLIFMIYKQKPRYERWKIIEATN--GNNYTFIDPTQLPYNEKWE 578
QY 606 FPRENLEFGKVLGSGARVWNAATAYGISTGVSIQVAVKMKKEKADSCEKALMSLKM 665
Db 579 FPRDKLGLKTLGAGAFGVKVEATAYGLGKEDNITRVAVKMLKASAHDPEREALMSLKI 638
QY 666 MTHLGHNDNINLLGACTLSPVVLIEFYCCYCGDLLNLYRSKREKPF----- 711
Db 639 LSHLQGHKNINLLGACTHGGPVLVITEYCHGDLNLFIRSKAENFLNFMVITPNFPEPV 698
QY 712 -----HRTW-----TEIFKEHNFSSYPTFOAHNSNMPGSREVQLHPPLDQLSGF 756
Db 699 TDYKNVSTERMFVRSDSGISSCTSDHYLDMPVTSRPTNSALDPSDDCQ----- 747

```

757 NGNSIHSEDEIEYENQKRLAEEEDLNVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLA 816  
 748 -----EWSWFLDMDLLRFSQVAGQLDFLAAKNCIHRDVA 783  
 817 ARNVLTGHKVKICDFGLARDILSDSSYYVRGNARLPVKWMAPESLFEGYITIKSDVWS 876  
 784 ARNVLTGSRVAKICDFGLARDIMNDSNYVVGKGNARLPVKWMAPESEFECVYTVQSDVWS 843  
 877 YGILLWEIFSLGVPYGGIPVDANFYKLIQSGFKMEQPPFYATSGIYFVMSCHAFPSRKR 936  
 844 YGIMLWEIFSLGKSPYNILVDSKFYKMKICGYQMSRDPFAPPMTYIMKMCWLEAER 903  
 937 PSFPMNLTSFLGQALAE-ABEACIRTSIHLPKQAAPQ 972  
 904 PTFKSIQMIQMLGSETSQDQOEYKNIPSEAVEQ 940

RESULT 11  
 Q918N6  
 ID Q918N6 PRELIMINARY; PRT; 977 AA.  
 AC Q918N6;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Fms.  
 GS CSFIR OR FMS.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN 1  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AB;  
 RX MEDLINE=203231166; PubMed=10862741;  
 RA Parichy D.M., Ransom D.G., Paw B., Zon L.I., Johnson S.L.;  
 RT "An orthologue of the kit-related gene fms is required for development  
 RT of neural crest-derived xanthophores and a subpopulation of adult  
 RT melanocytes in the zebrafish, Danio rerio.";  
 RL Development 127:3031-3044(2000).  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
 CC PROTEIN KINASES  
 DR EMBL; AF240639; AAF76872.1; --  
 DR HSSP; P11362; IFGK.  
 DR ZFIN; ZDB-GENE-001205-1; csflr.  
 DR InterPro; IPR007110; IG-like.  
 DR InterPro; IPR003598; IG\_c2.  
 DR InterPro; IPR003006; IG\_MHC.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR001824; RTKinaseII.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00047; Ig; 4.  
 DR Pfam; PF00069; pkinase; 1.  
 DR ProDom; PD000001; Prot\_kinase; 2.  
 DR SMART; SM00408; IGC2; 1.  
 DR SMART; SM00219; TyrKc; 1.  
 DR PROSITE; PS00835; IG LIKE; 3.  
 DR PROSITE; PS0107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
 DR ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;  
 KW Phosphorylation; Receptor; Transferase; Transmembrane;  
 KW Tyrosine-protein kinase.  
 FT VARIANT 238 238 N -> K.  
 FT VARIANT 615 615 V -> M.  
 SQ SEQUENCE 977 AA; 110187 MW; C91A2F339E746A58 CRC64;

Query Match 23.1%; Score 1214; DB 13; Length 977;  
 Best Local Similarity 33.4%; Pred. No. 2.5e-97;  
 Matches 329; Conservative 157; Mismatches 354; Indels 146; Gaps 34;

QY 68 PR-RQSEGTVYEATVEVABSGSITLQVQATPDGLSCLW---VFKHSLGCGPHFDLQN 123  
 DB 22 PRRLSSGAL--AGTDVILESGS---PLQLVCEGSGVTFELPRLAKH-----KRYISKEV 71  
 QY 124 RGVSMAILNVTETQAGEY-LLHIQSEGRANYTVLFTVVRDQ-LYV-----LRPFRK 176  
 DB 72 GKRSFRVEKTVDFGTGKCYKVMNGNDSLSNVHVFVDRSRVLFVSPSTSLR--YVRK 129  
 QY 177 MENQDALL-CISGVPPEPTVEWVLCSSHRSCKEEGPAVVRKEEK-----VLHFLGTDI 230  
 DB 130 -EGEDLLPLCL---LTDPEATDFTFRMDNGSAAPYGMNITDPRKGVLRNVHFGPNADY 185  
 QY 231 RCCARNALGRECTKLTIDLNQAPSTLPOLFJK-----VGEPLWIRCKAIHVRHGF 283  
 DB 186 ICCARIGGAERKVSIFSIINIIQRLRFP-PVYLKRNVEYKLVGERLQISCTTNNPFPYN 244  
 QY 284 LTWELEDKALESGSYFEMSTYTNRTMIRILLAFVSVSGRNDTGYTCSSSKIPSQSALV 343  
 DB 245 VTWTHSSRMPLKAE--EKSTMEGDLRAISILT-IPSVQLSTGNITCTGQNEAGANSST 301  
 QY 344 T---ILEKGFINAT---SSQEEY-----EIDPYEKFCFSVRFKAYPRIC---TWFSQ 388  
 DB 302 TQLLVVEEPIRLSPKLSKLTGRGLSIEVSEGGDVLGLVLEIAYPLTSHKWTPTSHN 361  
 QY 389 ASFPCEQR--GLEGDYSISKFCDHKN--RPGEYIFYAENDDAQFTKMTLNRKPKQVLA 444  
 DB 362 ASLP-ENRPFNNDRYEALLKRLNFEIQTGLVNVKSMKSASITFDIKMTKPVARV 420  
 QY 445 N-ASASQASCSDGYPPLPWTWKCKSDKSPNCTEEI-----PEGVWKKANRKYVFGQWSS 499  
 DB 421 KWENVTTLSRSGYGPAPSIILWYQCTGIRTTCPENTTDLQPIQTQVFQKESFGAVGE 480  
 QY 500 STLNMSEAGKGLLVKCCAYNSMGTSCTETFLNSPGFPFIQDNIISFYATIGLCLPFI-- 557  
 DB 481 SVLTVG-PNRMTVVCVAFNLVQGSDDTFSME-----VSDQIFTSAMCGSTVAMVVG 532  
 QY 558 VLVILCHIKYKQPRYESQLQMIQVTGPDNDNRYFYDFRDYEDLKWEPPRENLEFGKVL 617  
 DB 533 LLLIFMIYKQKPRYEIRWKIIEATN--GNNYTFIDPTQLPNEKWEPRDKLKGKTL 590  
 QY 618 GSGAGRNVNATAYGISKTGVSIOVAVKMLKEKADSCKEALMSELKMTHLGHHDNIVN 677  
 DB 591 GAGAFKVVETAYGLGKEDNITRVAVKMLKASHPDEREALMSELKLSHLGQHKNIYN 650  
 QY 678 LIGACTLSGPVYLIFEYCYGDLNVLNRSKRKF-----HRTW 715  
 DB 651 LIGACTHGGPVLIVITEYCHGDLNPLRSKAENFLNFMVTIPNFPPEPMTDYKNVSTERMF 710  
 QY 716 -----TEIPKEHNFSSYPTFOAHSNSMPSGSRVQLHPPLDQLSGFNGNSHSEDEIE 768  
 DB 711 VRSDSGISSTCSDHYLDMRPVTSRPTNSALDSSSECQ----- 747  
 QY 769 YENQKRLAEEEDLNVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVV 828  
 DB 748 -----EDSWPLDMDLLRFSQVAGQLDFLAAKNCIHRDVAARNVLLTNSRVA 795  
 QY 829 KICDFGLARDILSDSSYYVRGNARLPVKWMAPESEFEGYITIKSDVWSYGILLWEIFSG 888  
 DB 796 KICDFGLARDIMNDSNYVVGKGNARLPVKWMAPESEFECVYTVQSDVWSYGIMLWEIFSG 855  
 QY 889 VNPYGPVDANFYKLIQSGFKMEQPPFYATSGIYFVMSCHAFPSRKNPFLTSFLCC 948  
 DB 856 KSPYFNILVDSKFYKMKICGYQMSRDPFAPPMTYIMKMCWLNDAERPTFSKISQMIQR 915  
 QY 949 QLAE-AEEACIRTSIHLPKQAAPQ 973  
 DB 916 MLGETSQDQOEYKNIPTEAEAEQ 941

PRELIMINARY; PRT; 979 AA.

RESULT 12  
 Q8WN23  
 ID Q8WN23

```

AC Q8WN23;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE C-KIT.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Zemde D., Yuzbasiyan-Gurkan V.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES
DR EMBL; AF448148; AAL40833.1; -.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001824; RtkinaseIII.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00047; Ig; 2.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 2.
DR SMART; SM00409; IG; 3.
DR SMART; SM00219; Tyrcg; 1.
DR PROSITE; PS00835; IG_LIKE; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TVR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
DR ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW Phosphorylation; Receptor; Transmembrane;
KW Tyrosine-protein kinase.
SQ SEQUENCE 979 AA; 109753 MW; 46C30D5DBE8E33D3 CRC64;

Query Match 23.0%; Score 1209.5; DB 6; Length 979;
Best Local Similarity 31.3%; Pred. No. 6.1e-97;
Matches 332; Conservative 165; Mismatches 369; Indels 195; Gaps 35;

Qy 14 LLVLSVMILETVTNQDLVPKICVLISHENNGSSAGKSSYRMVRGSPEDLQCTPRQSE 73
Dy 11 LCVLILLLLGVTGSSQSPV-----SPGEPSSL----- 38
Qy 74 GTVYEAATVEVAESGSIITLOVLATPGDLSCULWFKHSLGCGPHDQNRGIVSMAILN 133
Dy 39 PSIIHPAKSELIVSGD-ELRLSCTDPGFVK--WTPE--TLG-QLNENTHNEWITEKA--- 89
Qy 134 VTEQAGEYLLHIQSERANYTVLFTVVRD-TQLYVLRBPYFRKMENQDALLCISEGVPE 192
Dy 90 ---EAGHTGNYTCNRDGLSRISYVVRDPAKFLVLDLPYKGEQNDTLVRC----- 138
Qy 193 PTVEWVLCSSHRESCKEPPAVRKEEVKHLFGTDIR-----C--CARNALGRE 241
Dy 139 PLTDEPVNYSLRGC--EGKPLPKDLTFVADPKAGITIRNVREYHRLCLHCSADOKGT 196
Qy 242 -CTKLFTDLNQAPOS-----TLPLFLKVGEPILWIRCKAIVHNGFGLTWELEDKALE 294
Dy 197 VLSKKEFTLKVAIRAIVPVVSVSKTSSLLKEGEAFSVMCFIKDVSSFVDSMWIKNSQQT 256
Qy 295 EGSYEMSTYSNRTMIRILLAFVSVGENDTYTCSSSKHPSQSALVTILE---KGFI 351
Dy 257 NAQTOSNWHHGDGFNEROEKLIISARVNDSGVFCYANNTFGSANVTITLLEWDKGF 316
Qy 352 NA-TSSQBEYEIDPEKFCFSFRKAYPR-IRCTWIFSOASF--PCEQRGLDEGYISKF 407
Dy 317 NIFPMMSITIFVNDGENDVLIYEYAYPKPEHQMIYNNRTTDKWDYPKSDNESNRY 376
Qy 408 CDH-----KNKPGEYIFYAENDDAQFTKMTFLNIRKPKQVLANASQA--SCSSDGY 458
Dy 377 YSELHLTRLKGEGGTYTFQVNSDVSNSSTFNVYVNTKPEILTHESLTNGMLQCWVAG 436

```

```

Qy 459 PLPSWTWKCKSDKSPNCTEEI--PEGVWKNKANRVKFGQWVSSSTLMSBAGKLLVKCCA 517
Dy 437 PEPAVDWYFCPGAEOQRCVPIGPMVQVQNSLSLSPGKLVQSSIDYSAPKNGTVECR 496
Qy 518 YNSMTSCETIFLNSPGPPFIQDN-----ISFYATIGLCPLFFIVLVILVI 563
Dy 497 YNVVGRS--SAFFN---PAFKGSKGEQHPHTLFTPLLLIGFVIAAGM---MCIIVMIL 546
Qy 564 CHYKKQFRYESOLOMI-QVTGPLDNEYFYVDFRDEYDYLKWEFFRENLEFGKVLGSGAF 622
Dy 547 TYKLOKPMYEVQWKVBEING---NNVYVIDPTQLPYDHKWEFPNRLSFGKTLGAGAF 603
Qy 623 GRVMNATAYGISTGVISQVAVKMLKEKADSCKEALMSELKMTLGHHDHINVLIGAC 682
Dy 604 GKVEATAYGLIKSDAAMTVAVKMLKPSAHLTEREALMSELKVLVSLGNHMINVLIGAC 663
Qy 683 TLSGPVYLIFEYCCYGDLLNLYLRSKREKPHRTWTWTFIKKHNFSSYPTFOAHNSMPGSR 742
Dy 664 TVGGPTLVITEYCCYGDLLNFLRRKDSFICSKQE---DHG----- 701
Qy 743 EVQLHPPLDQLSGFNGNSIHSDEI-----EVENOK-----RLAEEE 779
Dy 702 EVALYK-----NLLHSESSCSNSTEYMDMKPGCVSVVPTKADKRRSARIGSYI 751
Qy 780 EEDLN-----VLTPEDLLCFAYQVAKGMEFLFEKSCVHRDLAARNVLVTHGKVKI 830
Dy 752 EEDVTPAIMEDDELALDLEDLLSFSYQVAKGNAFLASKNCIHRDLAARNILLTHGRITKI 811
Qy 831 CDFGLARDILSDSYVVRGNARLPVKWMAPELSFEGITYTIKSDVMSYGILLWEIFSLGVN 890
Dy 812 CDFGLARDIKDSDNYVVRGNARLPVKWMAPELSFEGITYTIKSDVMSYGIFLWELFSLGSS 871
Qy 891 PYRGIPVDANFYKLIQSCKFMEQPFYATYEGIVFMOSWAFPSRKEPSPNLTSLFLGCOL 950
Dy 872 PYPGMPVDSKFYKMIKEGRFMLSPHAPAEYDIMKTCMDADPLKRPKTFQIVOLLEKQI 931
Qy 951 AEAEAEACIRTSIHLPKQAAPQORGGGLRAQSPQORQVKIHRER 991
Dy 932 SDSTN-----HIYSNLAN-----CSPNPERPVVDHSVR 959

RESULT 13
O97744
ID O97744; PRELIMINARY; PRT; 964 AA.
AC O97744;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Mast/stem cell growth factor receptor (Fragment).
GN KIT1*0101.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hampshire breed; TISSUE=Leukocyte;
RX MEDLINE=98391767; PubMed=9724328;
RA Marklund S., Kiljas J., Rodriguez-Martinez H., Roennstrand L., Funa K.,
RA Moller M., Edfors-Lilja I., Andersson L.;
RT "Molecular evolution of the dominant white phenotype in the domestic
RT pig.";
RL Genome Res. 8:826-833 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Hampshire breed; TISSUE=Leukocyte;
RA Marklund S.;
RL Thesis (1997), Department of Animal Breeding and Genetics.
RL Swedish University of Agricultural Sciences, Uppsala, SWEDEN.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.

```

DR EMBL; AJ223228; CAAL1196.1; --  
 DR HSP; P11362; 1FGK.  
 DR InterPro; IPR003599; IG.  
 DR InterPro; IPR007110; IG-like.  
 DR InterPro; IPR003006; IG MHC.  
 DR InterPro; IPR000719; Prot kinase.  
 DR InterPro; IPR001824; RTKinaselII.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR Pfam; PF00047; ig; 1.  
 DR Pfam; PF00069; pkinase; 1.  
 DR ProDom; PD000001; Prot\_kinase; 2.  
 DR SMART; SM00409; IG; 3.  
 DR SMART; SM00219; Tyrc; 1.  
 DR PROSITE; PS00835; IG-Like; 3.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.  
 DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
 KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;  
 KW Phosphorylation; Receptor; Transferase; Transmembrane;  
 KW Tyrosine-protein kinase.  
 FT NON TER 964 964  
 SQ SEQUENCE 964 AA; 108287 MW; 90B72EDFAB1358B7 CRC64;

Query Match 22.9%; Score 1208; DB 6; Length 964;  
 Best Local Similarity 32.6%; Pred. No. 8.1e-97;  
 Matches 319; Conservative 168; Mismatches 362; Indels 130; Gaps 32;

QY 38 LISHENNGSSAGKSSYRVVRGSPEDLOCTPRQSGTVYEAATVEASGSLTQVOLA 97  
 Db 16 LLRLRVQTGSS--QPSV-----SPEL-----SPSPHPAKSELIVSAGD--EIRLFT 59  
 QY 98 TPGDLSCLVWVFXHSLGCPHDLQNRGIVSMALLNVTTQAGEYLLHQTSERA----- 151  
 Db 60 DFG--SVKWTPE--TLG-----OLSENTHAEWIV-----EKAEAMNTG 93  
 QY 152 NYTVL-----FTVNRVDTQ--LYVLRPFRKMNODAL-----LCISE 188  
 Db 94 NYTCTNEGGLSSIVYFVRDPEKFLVDPPLYGK--EDNALVRCPLTDEVTNYSLTGE 152  
 QY 189 GVPEP--TVEWLCSSHRESCKEAGFAVRKEKVLHFGTDIRC--CARNALGRE--CTK 244  
 Db 153 GKPLPKDITFV-----ADPKAGITIRNVKREYHRL-----CLHSANQGKSVLSK 198  
 QY 245 LFTIDLNAPOS-----TLPLQLKVGCEPLWIRCKAIVNHGFLTWELEKALEBEGY 298  
 Db 199 KFTLVRAAIAFVPVAVSKASYLLREGEFAVMCLIKDVSSVDSMW--IRENSQTKAQV 257  
 QY 299 FEMTSTYNTRMIRILLAFVSSVGRNDTGYYTCCSSKHPQSOSALVTILE---KGFINA-T 354  
 Db 258 KRNSWHQGDFFLQERLTISSARVNDGVPVPCYANNTFGSANVTITLVDKGFINIFP 317  
 QY 355 SSQEEYEDIPYKCFVRFKAYPR--IRCTWIFSQ--ASFPCEQRGLBEGYSISKFCDHK 411  
 Db 318 MNMTVTFVNDGEDVDLVEYEAYPEKPHRQWIMYRNTATDKWEDYPKSENSNIRYSEL 377  
 QY 412 N-----KPEYIFYAENDDAQTKMFTLNIRKKPQVILAN--ASASQASCSGDPPLPS 462  
 Db 378 HLTRFLKTEGGTYTFLVSNADVNSVTFNVVNTKPEILTHDRLNMGMLQCVAAAGFPEPT 437  
 QY 463 WTKKCKSDKSPNCTEEI--PEGVMNKNKRVFGQWVSSSTLNMSEAGLLAVKCCAYNSM 521  
 Db 438 IDWYFCPTGTCRQSVVPVGVQVQVQVSSVSPFGKLVHSSIDYSAFKHNGTVECRANDV 497  
 QY 522 GTSCEITFLNSPGPPFIQNIISFYATIGLCLPFIV-----VLIVLCHKYKQPRYES 575  
 Db 498 GKS--SAFFN-----FAPKEQIHATLFTPLLIGFVIAAGMCIIMVILTYKLOKPYEV 551  
 QY 576 QLQMI--QVTPGLDNEYVDYFRYEDYDLKWEFPRENLEFGKVLGSGAFGRVNNATAYGIS 634  
 Db 552 QWKVBEING-----NNYVIDPTQPDYDKHWEFPRENRLSFGKTLGAGAFGKVVATEAYGLI 608  
 QY 635 KTVGSVQVAVKMLKERADSKCEKALMSKMTWHLGHHDHINVLGLACTLSGPPYLIFEY 694

Db 609 KSDAATVAVKMLKPSAHLTEREALMSKLVSYLGNMNVNLLGACTIGGFTLVITY 668  
 QY 695 CCYGDLLNYLRSKREKFRHTWTEIFKEHNFSSYPTFOAHSSNMPGSGREVQLHPPLDOLS 754  
 Db 669 CCYGDLLNFLRRKRDSTFCSKQEDHAAALYKLLHKSCESSSTNMYMDKPGVSYV 728  
 QY 755 GFNGNSIHSDEIYENOKRLAEEEDLNVLTPEDLCLCFAYQVAKMEFLFKSCVHRD 814  
 Db 729 PTKADKRSARIGSYIERDVTPTAIMEDELALDLEDLLSFSYQVAKGMAFLASKNCIHRD 788  
 QY 815 LAARNVLVTHGKVKVICDFGLARDILSDSSVYVVGNAFLPVKMAPELPGIYTIKSDV 874  
 Db 789 LAARNVLLTGRITKICDFGLARDIKDNSYVVGNAFLPVKMAPELPGIYTIKSDV 848  
 QY 875 WSYGILLWEIFSLGVNPPYGPVVDANFYKLIQSFQKMEQPPYATEGIVFVQMSQWAPDSR 934  
 Db 849 WSYGIFLWELFSLGSSPYPCMPVDSKPYKIKESGFRMLSPHEHAPAEYMDINKTCWDADPL 908  
 QY 935 KRPSFPNLTSFLGCOLA 953  
 Db 909 KRPTFKIVQLIEKQISES 927

RESULT 14  
 Q9TQ01 PRELIMINARY; PRT; 964 AA.

AC Q9TQ01;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Mast/stem cell growth factor receptor (Fragment).  
 GN KITI\*0201.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxId=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Yorkshire / Landrace cross; TISSUE=Leukocyte;  
 RA Marklund S.;  
 RL Thesis (1997), Department of Animal Breeding and Genetics,  
 RL Swedish University of Agricultural Sciences, Uppsala, SWEDEN.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Yorkshire / Landrace cross; TISSUE=Leukocyte;  
 RX MEDLINE=98391767; Pubmed=9724328;  
 RA Marklund S., Kijas J., Rodriguez-Martinez H., Roennstrand L., Funa K.,  
 RA Moller M., Edfors-Lilja I., Andersson L.;  
 RT "Molecular evolution of the dominant white phenotype in the domestic  
 RT pig.";  
 RL Genome Res. 8:826-833(1998).  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
 CC PROTEIN KINASES.  
 DR EMBL; AJ223228; CAAL1197.1; --  
 DR HSP; P11362; 1FGK.  
 DR InterPro; IPR003599; IG.  
 DR InterPro; IPR007110; IG-like.  
 DR InterPro; IPR003006; IG MHC.  
 DR InterPro; IPR000719; Prot kinase.  
 DR InterPro; IPR001824; RTKinaselII.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR Pfam; PF00047; ig; 1.  
 DR Pfam; PF00069; pkinase; 1.  
 DR ProDom; PD000001; Prot\_kinase; 2.  
 DR SMART; SM00409; IG; 3.  
 DR SMART; SM00219; Tyrc; 1.  
 DR PROSITE; PS00835; IG-Like; 3.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
 DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.  
 DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.

KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;  
KW Phosphorylation; Receptor; Transferase; Transmembrane;  
KW Tyrosine-protein kinase.  
FT NON TER 964 964  
SQ SEQUENCE 964 AA; 108315 MW; 996C3C46201358A8 CRC64;

Query Match 22.9%; Score 1207; DB 6; Length 964;  
Best Local Similarity 32.6%; Pred. No. 1e-96;  
Matches 319; Conservative 168; Mismatches 362; Indels 130; Gaps 32;

QY 38 LI SHENNAGSSAGKPSYRMVRGSPEDLOCTPRROEGTVYEAATVEVAESGSLTLQVOLA 97  
DB 16 LLLRVQTGSS--QPSV-----SPEL-----SPPSIHPAKSELIVSAGD-EIRLFCT 59  
QY 98 TPGDLSCLVWFVFKHSLGQCPHFQDLQNRGIVSMALNVTTQAGEYLLHIQSERAAAA 151  
DB 60 DPG--SVKWTFF--TLG-----QLSENTHAEWIV---EKAEMNTG 93  
QY 152 NYTVL-----FTVNVDTQ-LVLRPRPRKVENQDAL-----LCISE 188  
DB 94 NYTCTNEGGLSSSIYVVRDPKFLVDPPLPYGK-EDNDALVRCPLTDPVTNYSLTQCE 152  
QY 189 GVPEP-TVEWVLCSSHRESKEEGPAVVRKEKVLHFLGTDIRC--CARNALGRE-CTK 244  
DB 153 GKPLPKDITFV-----ADPKAGITIRNVKREYHRL-----CLHCSANOGGKSVLSK 198  
QY 245 LFTIDLNAPOS-----TLPOLFLKVGEPWIRCKAIHVNHGFLGTWELEDKALEEGSY 298  
DB 199 KFTLVKRAIRAVPVVAVSKASYLLREGPEFAMVCLIKDVSSVDSMW-IRENSQTKAQV 257  
QY 299 FEMSTYSNRTMIRILLAFVSVGNRDTGYTCTSSSKHPSQSAVLTILE---KGFINA-T 354  
DB 258 KRNSWHQGDFFLROERLTISSARVNDSGVFMCYANNTFGSANVTTTLEVDVKGFINFP 317  
QY 355 SSOEYEDIDPKCFKCFVRFKAYPR-IRCTWLFSQ--ASFPCQORGLDGVYSIFPCDHK 411  
DB 318 MNMTTVFVNDGDVLDLVEYEAIPKPEHRQWIMNRTADKWDYKPKENESNIIRYSEL 377  
QY 412 N-----KPGEYIFVANDDAQFTKMTFLNIRKKPOVLAN--ASASQASCSDDGYPLPS 462  
DB 378 HLTRLKGTGEGTYTFLVSNADVSSVTFVNVYNTKPEILTHDLNGLMLQCVAAGFPPT 437  
QY 463 WTKKCKSKSPNCTEEL-PEGVWKNKANRKYGVQWSSVSTLWNSAGKGLLVKCCAYNSM 521  
DB 438 IDWYFCPGTEQRCSVPVGPVDVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 497  
QY 522 GTSCTEILNSGPPFPFQDNISFATIGLCLPFIV-----VLIVLCHKYKKQFRYES 575  
DB 498 GKS--SAFFN-----FAFKEQIHAHTLFTPLLLGFVIAAGMMCIIVMLTYKYLQKPMYEV 551  
QY 576 QLOMI-QVTGPLDNEYFYVDRDYDLKWEPPRENLEFPGKVLGSGAFGRVNMATAYGIS 634  
DB 552 QWKVVEENG---NNYVIDPTQLPYDHKWEPPRNLSFGKTLGAGAFKVVVEATAYGLI 608  
QY 635 KTVGSIQVAVKMLKPKADSCKEALMSLKMTHLGHONIVNLGACTSLSGPVYLIFPY 694  
DB 609 KSDAAMTAVKMLKPSAHLTEREALMSLKVLSYLGHNHNVNLLGACTIGGPTLVITEY 668  
QY 695 CCYGBLLNVLRSKREKFRHTWTETFEKHNFSYPTFOAHNSNMPGSRREVQLHPDQLS 754  
DB 669 CCYGBLLNVLRSKREKFRHTWTETFEKHNFSYPTFOAHNSNMPGSRREVQLHPDQLS 728  
QY 755 GFNGNSIHSEDEIEYENOKLAEEREDLVLTFFEDLLCFAYQVAKGMEFLFKSVCHRD 814  
DB 729 PTKAKRARSATGSIYERDVTVAIMEDDELALDLELLSFSQVAKGMAFLASKNCIHRD 788  
QY 815 LAARNVLVTHGKVVKICDFGLARDILSDSYVYVGNARLPVKNMPPESFEGITYIKSDV 874  
DB 789 LAARNVLLTHGRITTKICDFGLARDIKNDVYVYVGNARLPVKNMPPESFEGITYIKSDV 848  
QY 875 WSYGILLWEIFSLGVNYPGIPVDANFYKLIQSGFKQEPFVATEGIIYFMQSCWAFDSR 934  
DB 849 WSYGIFLWELFSLGSSPYFGMPVDVDSKFKYKMKEGFRLSPHAPVEMYDIMKTCWADFL 908

QY 935 KRSPFNLTSLFGCOLAEEA 953  
DB 909 KRPTFKOIVQVLEIKOISES 927

RESULT 15

Q9TQ00 PRELIMINARY; PRT; 964 AA.  
AC Q9TQ00;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Mast/stem cell growth factor receptor (Fragment).  
GN KITI\*0202.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Yorkshire / Landrace cross; TISSUE=Leukocyte;  
RA Marklund S.;  
RL Theis (1997), Department of Animal Breeding and Genetics,  
RL Swedish University of Agricultural Sciences, Uppsala, SWEDEN.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Yorkshire / Landrace cross; TISSUE=Leukocyte;  
RX MEDLINE=98391767; PubMed=9724328;  
RA Marklund S., Kijas J., Rodriguez-Martinez H., Roennstrand L., Funa K.,  
RA Moller M., Edfors-Lilja I., Andersson L.;  
RT "Molecular evolution of the dominant white phenotype in the domestic pig";  
RL Genome Res. 8:826-833(1998).  
CC - - SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).  
CC - - SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
CC PROTEIN KINASES.  
DR EMBL; AJ223230; CAAL1198.1; - -  
DR HSP; P11362; IFGK.  
DR InterPro; IPR003599; IG.  
DR InterPro; IPR007110; IG-like.  
DR InterPro; IPR003006; IG\_MHC.  
DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR001824; RtkinaseII.  
DR Pfam; PF00047; Ig; 1.  
DR Pfam; PF00069; pkinase; 1.  
DR ProDom; PD000001; Prot\_kinase; 2.  
DR SMART; SM00409; IG; 3.  
DR SMART; SM00219; Tyrc; 1.  
DR PROSITE; PS0835; IG LIKE; 3.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00101; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;  
KW Phosphorylation; Receptor; Transferase; Transmembrane;  
KW Tyrosine-protein kinase.  
FT NON TER 964 964  
SQ SEQUENCE 964 AA; 108287 MW; 1B21292A962E9191 CRC64;

Query Match 22.9%; Score 1207; DB 6; Length 964;  
Best Local Similarity 32.6%; Pred. No. 1e-96;  
Matches 319; Conservative 168; Mismatches 362; Indels 130; Gaps 32;

QY 38 LI SHENNAGSSAGKPSYRMVRGSPEDLOCTPRROEGTVYEAATVEVAESGSLTLQVOLA 97  
DB 16 LLLRVQTGSS--QPSV-----SPEL-----SPPSIHPAKSELIVSAGD-EIRLFCT 59  
QY 98 TPGDLSCLVWFVFKHSLGQCPHFQDLQNRGIVSMALNVTTQAGEYLLHIQSERAAAA 151  
DB 60 DPG--SVKWTFF--TLG-----QLSENTHAEWIV---EKAEMNTG 93







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2003, 07:13:34 ; Search time 62.4685 seconds

(without alignments)

2520.578 Million cell updates/sec

Title: US-09-919-408A-2

Perfect score: 5264

Sequence: 1 MRALQSRDRRLLLVLSV.....RGGLRAQSPQVKIHRS 992

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5264	100.0	992	14	Murine flk-2. Mus
2	5264	100.0	992	16	Flk2 receptor prot
3	5264	100.0	992	16	Mouse flk-2. Mus
4	5264	100.0	992	17	Human foetal liver
5	5264	100.0	992	18	Murine flk-2 recep
6	5264	100.0	992	20	Murine flk-2 prote
7	5264	100.0	992	23	Mouse receptor pr
8	5264	100.0	992	23	Murine receptor pr
9	5266	99.8	992	14	Murine flk-2 recep

10	5256	99.8	992	14	AA331375	Murine flk-2. Mus
11	5239	99.5	992	13	AA280318	Murine flk-2. Mus
12	5102	96.9	1000	16	AA81868	Flk2/flt3 tyrosine
13	4436.5	84.3	993	16	AA75961	Human STK-1. Homo
14	4429.5	84.1	993	16	AA67816	Flk2 receptor prot
15	4429.5	84.1	993	16	AA67536	Human flk-2. Homo
16	4429.5	84.1	993	17	AA97419	Murine foetal live
17	4429.5	84.1	993	18	AA19873	Human flk-2 recep
18	4429.5	84.1	993	20	AA08617	Human flk-2 protei
19	4429.5	84.1	993	23	AB70916	Human receptor pro
20	4429.5	84.1	993	23	AA25819	Human receptor pro
21	4426.5	84.1	993	14	AA37503	Human flk-2. Homo
22	4421.5	84.0	993	14	AA44995	Human flk-2 recep
23	4420.5	84.0	993	16	AA81869	Human Flk2/flc3 ty
24	4417.5	83.9	1167	14	AA31376	Human flk-2. Homo
25	4409.5	83.8	983	19	AA63588	Human receptor typ
26	4408	83.7	986	19	AA63587	Human receptor typ
27	4408	83.7	986	19	AA63589	Human receptor typ
28	4404	83.7	994	19	AA63586	Human receptor typ
29	2806.5	53.3	665	19	AA63585	Human receptor typ
30	2373.5	45.1	749	19	AA78002	Protein pMON32390.
31	2326.5	42.3	481	15	AA47579	Soluble Flk-2. Mu
32	1461	27.8	307	23	AB81198	Human FLT3 protein
33	1223.5	23.2	972	23	AA11935	Colony stimulating
34	1223	23.2	975	22	AAE07144	Murine Kit/stem ce
35	1223	23.2	975	22	AAE07148	Mutant murine Kit/
36	1221.5	23.2	972	23	AAU79039	Human macrophage c
37	1221.5	23.2	972	23	AAU11936	Colony stimulating
38	1221.5	23.2	972	23	AAU11941	Colony stimulating
39	1221.5	23.2	972	23	AAU73585	Colony stimulating
40	1221.5	23.2	1055	22	ABG15479	Novel human diagno
41	1220.5	23.2	972	23	AAU11934	Colony stimulating
42	1220.5	23.2	972	23	AAU11937	Colony stimulating
43	1219.5	23.2	972	23	AAU11940	Colony stimulating
44	1216.5	23.1	972	23	AAU11938	Colony stimulating
45	1215	23.1	977	21	AAV51322	Bovine c-Kit bK-1

## ALIGNMENTS

RESULT 1  
AA337502  
ID AA337502 standard; Protein; 992 AA.

AA337502;  
AC  
XX  
DT 25-MAR-2003 (updated)  
DT 19-OCT-1993 (first entry)  
XX  
XX  
DE Murine flk-2.

Murine; receptor; protein; tyrosine kinase; pTK; flk-2; primitive;  
hematopoietic cell; mature; family; conserved; region;  
catalytic domain; c-kit; fetal liver kinase; flk; fetal; spleen;  
thymus; adult; brain; bone marrow; multipotential; CFU-Blast colony;  
hierarchy; transduction; T-lymphoid; lineage.

Mus musculus.

Key	Location/Qualifiers
Peptide	1..27
Protein	/note= "Hydrophobic leader sequence"
Domain	28..992
Region	/note= "Mature murine flk-2"
Domain	28..544
Region	/note= "Extracellular receptor domain"
Domain	545..564
Domain	/note= "Transmembrane region"
Domain	565..992
Domain	/note= "Intracellular catalytic domain"
Domain	618..623
Domain	/note= "Catalytic sub-domain"

FT Domain 811..819 /note= "Catalytic sub-domain"  
 FT Domain 832..834 /note= "Catalytic sub-domain"  
 FT Domain 857..862 /note= "Catalytic sub-domain"  
 FT Domain 872..878 /note= "Catalytic sub-domain"  
 FT Region 736..812 /note= "Catalytic sub-domain"  
 FT /note= "flk-2 signature sequence"  
 XX w09310136-A1.  
 XX 27-MAY-1993.  
 XX 16-NOV-1992; 92WO-US09893.  
 XX 15-NOV-1991; 91US-0793065.  
 XX (UYPR-) UNIV PRINCETON.  
 XX Lemischka IR;  
 XX WPI; 1993-182479/22.  
 XX N-PSDB; AAQ40914.  
 XX Totipotent haematopoietic stem cell receptors, their ligands and  
 PT DNA sequences - for treating anaemia(s) and bone marrow damage  
 PT due to e.g. cancer chemotherapy or radiotherapy  
 XX Claim 37; Fig 1a; 127pp; English.  
 XX This sequence represents the murine receptor protein tyrosine kinase  
 CC (PTK), flk-2. The nucleic acid encoding this receptor is expressed  
 CC in primitive hematopoietic cells and not in mature hematopoietic  
 CC cells. Members of this family of PTK's can be recognised by the  
 CC conserved amino acid regions in the catalytic domain. This family  
 CC of PTK's also contains c-kit. These new receptors are termed fetal  
 CC liver kinases (flk's) after the tissue in which they were discovered.  
 CC flk-2 is also expressed in fetal spleen, fetal thymus, adult brain  
 CC and adult bone marrow. flk-2 is expressed in individual multipotential  
 CC CFU-Blast colonies capable of generating numerous multilineage colonies  
 CC upon replating. It is likely therefore, that flk-2 is expressed in  
 CC the entire primitive portion of the hematopoietic hierarchy. This is  
 CC consistent with flk-2 being important in transducing putative self-  
 CC renewal signals from the environment. flk-2 is the first receptor  
 CC PTK known to be expressed in the T-lymphoid lineage.  
 CC (Updated on 25-WAR-2003 to correct PN field.)  
 XX Sequence 992 AA;  
 SQ Query Match 100.0%; Score 5264; DB 14; Length 992;  
 Best Local Similarity 100.0%; Pred No. 0;  
 Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MRALAQRSDRRLLLVLSVILETVTNQDLPVVKVLI SHENNGSAGKPSVYRMVRS 60  
 DB 1 MRALAQRSDRRLLLVLSVILETVTNQDLPVVKVLI SHENNGSAGKPSVYRMVRS 60  
 QY 61 PEDLOCTPRQSEGTVEAATVEAEGSITLQVLATPGDLSCLVFKHSSLCQCPHFD 120  
 DB 61 PEDLOCTPRQSEGTVEAATVEAEGSITLQVLATPGDLSCLVFKHSSLCQCPHFD 120  
 QY 121 LQNRGIVSMALNVTQAGYLLHIO SERANYTVLTVNVRDQLVLRARPYPRKMEQ 180  
 DB 121 LQNRGIVSMALNVTQAGYLLHIO SERANYTVLTVNVRDQLVLRARPYPRKMEQ 180  
 QY 181 DALLCISGVPETVVEVLCSSHRESCKE GPVVRKEEKVHLFGTDIRCCARNALGR 240  
 DB 181 DALLCISGVPETVVEVLCSSHRESCKE GPVVRKEEKVHLFGTDIRCCARNALGR 240  
 QY 241 ECTKLFITDINQAPQSTLPQLFLKVGSPWIRCKAIHVNHGFLTWELEDKALEEGSYFE 300  
 DB 241 ECTKLFITDINQAPQSTLPQLFLKVGSPWIRCKAIHVNHGFLTWELEDKALEEGSYFE 300

DB 241 ECTKLFITDINQAPQSTLPQLFLKVGSPWIRCKAIHVNHGFLTWELEDKALEEGSYFE 300  
 QY 301 MSTYSTNRTMIRILLAPVSSVGRNDGYTCSSSKHPSQSALVTILEKGFINATSSQEEY 360  
 DB 301 MSTYSTNRTMIRILLAPVSSVGRNDGYTCSSSKHPSQSALVTILEKGFINATSSQEEY 360  
 QY 361 EIDPYEKFCFSVRFKAYPRIRCTWIPFSAQSPCEORCLEGDYSISKFCDHKNKPGEIFY 420  
 DB 361 EIDPYEKFCFSVRFKAYPRIRCTWIPFSAQSPCEORCLEGDYSISKFCDHKNKPGEIFY 420  
 QY 421 AENDDAQFTKMFTLNIRKKPOVLANASASQACSSSDGYPLPSTWTKKCDKSPNCTBEIP 480  
 DB 421 AENDDAQFTKMFTLNIRKKPOVLANASASQACSSSDGYPLPSTWTKKCDKSPNCTBEIP 480  
 QY 481 EGVNKKANRKFVGQWVSSSTLANSEAGKLLVKCCAYNSMGTSCTEIFLNSPGPPFIQ 540  
 DB 481 EGVNKKANRKFVGQWVSSSTLANSEAGKLLVKCCAYNSMGTSCTEIFLNSPGPPFIQ 540  
 QY 541 DNISFYATIGLCPLPFIWLVILICHYKKQFRYESQLOMIQVGTGPDNEFYVDPRDYEY 600  
 DB 541 DNISFYATIGLCPLPFIWLVILICHYKKQFRYESQLOMIQVGTGPDNEFYVDPRDYEY 600  
 QY 601 DLKWEPPRENLEFGKVLGSGAFGRVNNATAYGISKTGVSIOQAVKMLKEKADSCKEKALM 660  
 DB 601 DLKWEPPRENLEFGKVLGSGAFGRVNNATAYGISKTGVSIOQAVKMLKEKADSCKEKALM 660  
 QY 661 SELKMTLHGHNDINVLGACTLSGPVYLIFEYCCYGDLLNLYLSKREKFRHTWTBIFK 720  
 DB 661 SELKMTLHGHNDINVLGACTLSGPVYLIFEYCCYGDLLNLYLSKREKFRHTWTBIFK 720  
 QY 721 EHNFSYPTFAHNSNSMPSGSRVQLHPDLQSLSGFNNGSIHSEDEIEYENQKLAEEEE 780  
 DB 721 EHNFSYPTFAHNSNSMPSGSRVQLHPDLQSLSGFNNGSIHSEDEIEYENQKLAEEEE 780  
 QY 781 EDNLNLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840  
 DB 781 EDNLNLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840  
 QY 841 SDSSYVVRGNARLPVKWMAPESLFEGITTIKSDVMSYGILLWEIFSLGVNPGIPVDAN 900  
 DB 841 SDSSYVVRGNARLPVKWMAPESLFEGITTIKSDVMSYGILLWEIFSLGVNPGIPVDAN 900  
 QY 901 FYKLIOGPFMEOPFVATEGIYFVMSQWAFDSRKRPSFNLTSFLGCQLAABEACIRT 960  
 DB 901 FYKLIOGPFMEOPFVATEGIYFVMSQWAFDSRKRPSFNLTSFLGCQLAABEACIRT 960  
 QY 961 SIHLPKQAPQQRGGLRAQSPQKVIHRERS 992  
 DB 961 SIHLPKQAPQQRGGLRAQSPQKVIHRERS 992  
 RESULT 2  
 AAR67815  
 ID AAR67815 standard; Protein; 992 AA.  
 XX AAR67815;  
 XX AC AC  
 XX 25-MAR-2003 (updated)  
 XX 16-AUG-1995 (first entry)  
 XX Flk2 receptor protein-tyrosine-kinase.  
 XX Mouse Flk2; receptor protein-tyrosine-kinase; primitive  
 KW hematopoietic cell; fetal liver kinase; diagnostic ligand  
 KW isolation; bone marrow disease therapy.  
 XX Mus musculus.  
 OS  
 XX Key Location/Qualifiers  
 FH 1..27  
 FT Peptide /note= "signal peptide"  
 FT Protein 28..992  
 FT /note= "mature protein"

FT Domain 28..544  
FT /note= "extracellular receptor domain"  
FT 545..564  
FT /note= "transmembrane region"  
FT 565..592  
FT /note= "intracellular catalytic domain"  
FT 618..623  
FT /note= "catalytic sub-domain"  
FT 811..819  
FT /note= "catalytic sub-domain"  
FT 832..834  
FT /note= "catalytic sub-domain"  
FT 857..862  
FT /note= "catalytic sub-domain"  
FT 872..878  
FT /note= "catalytic sub-domain"  
FT 709..785  
FT /note= "signature sequence typical of Flk2"  
XX WO9500554-A2.  
XX  
XX 05-JAN-1995.  
XX  
XX 17-JUN-1994; 94WO-US06944.  
XX  
XX 18-JUN-1993; 93US-0080244.  
XX 21-JUN-1993; 93US-0081508.  
XX 23-NOV-1993; 93US-0157490.  
XX  
XX (UYPR-) UNIV PRINCETON.  
XX  
XX Lemischka IR;  
XX  
XX WPI; 1995-052014/07.  
XX N-PSDB; AAQ81012.  
XX  
XX Ligand for receptor protein tyrosine kinase - useful for the  
XX stimulation of primitive haematopoietic stem cells causing  
XX proliferation and/or differentiation  
XX  
XX Disclosure; Fig 1a; 131pp; English.  
XX  
XX The sequence corresponds to a mouse Flk2 (fetal liver kinase)  
XX receptor protein-tyrosine-kinase, which is expressed in primitive  
XX hematopoietic cells but not in mature hematopoietic cells. The  
XX protein is useful in isolation of receptor ligands, which have  
XX applications in diagnosis of bone marrow disorders and in  
XX stimulating proliferation and/or differentiation of primitive  
XX hematopoietic stem cells.  
XX (Updated on 25-MAR-2003 to correct PN field.)  
XX  
XX Sequence 992 AA;

Query Match 100.0%; Score 5264; DB 16; Length 992;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALQSRDRRLLLVLSVMIETVTNODLPVVKVLSHENNGSSAGKPSYRMVRGS 60  
Db 1 MRALQSRDRRLLLVLSVMIETVTNODLPVVKVLSHENNGSSAGKPSYRMVRGS 60  
QY 61 PEDLOCTPRQSEGTVEAATVEAEGSITLQVQLATPGDLSCLWVFXHSLGCQPHD 120  
Db 61 PEDLOCTPRQSEGTVEAATVEAEGSITLQVQLATPGDLSCLWVFXHSLGCQPHD 120  
QY 121 LQNRGIVSMAILNVETQAGEYLLHQSERANYTLFTVNVVDTQYVLRFRKMNQ 180  
Db 121 LQNRGIVSMAILNVETQAGEYLLHQSERANYTLFTVNVVDTQYVLRFRKMNQ 180  
QY 181 DALLCISGVPEPTVEWVLCSSHRESCKEKGPAVVRKEEKVLHFLFGTDIRCARNALGR 240  
Db 181 DALLCISGVPEPTVEWVLCSSHRESCKEKGPAVVRKEEKVLHFLFGTDIRCARNALGR 240

QY 241 ECTKLTIDLNOAQSTLPOLFLKVGRLWIRCKAIHVNHGFLTWELDKALEBSYFE 300  
Db 241 ECTKLTIDLNOAQSTLPOLFLKVGRLWIRCKAIHVNHGFLTWELDKALEBSYFE 300  
QY 301 MSTYSTNRTWIRILLAFVSSVGRNDTGYTCCSSSKHPSQSALVTILEKGFINATSSQBEY 360  
Db 301 MSTYSTNRTWIRILLAFVSSVGRNDTGYTCCSSSKHPSQSALVTILEKGFINATSSQBEY 360  
QY 361 EIDPYEKFCPSVRPKAYPRIRCTWIFSQASPCCORGLEDGYSISKFCDDHKNKPGEYIFY 420  
Db 361 EIDPYEKFCPSVRPKAYPRIRCTWIFSQASPCCORGLEDGYSISKFCDDHKNKPGEYIFY 420  
QY 421 AENDDAQFTKMFNLINRKPKQVLANASASOASCSGDGVPPLPSWTWKCKSDKSPNCTEIP 480  
Db 421 AENDDAQFTKMFNLINRKPKQVLANASASOASCSGDGVPPLPSWTWKCKSDKSPNCTEIP 480  
QY 481 EGVNKKANRKFQGVQWVSSSTLNNSSEAGKLLVKCCAYNSMGTSCTEIFLNSPFPFPIQ 540  
Db 481 EGVNKKANRKFQGVQWVSSSTLNNSSEAGKLLVKCCAYNSMGTSCTEIFLNSPFPFPIQ 540  
QY 541 DNISFYATIGLCPLFFIWLIVLICHKYKKQFRYESQLOMIQVTGFLDNEYFYVDFRDY 600  
Db 541 DNISFYATIGLCPLFFIWLIVLICHKYKKQFRYESQLOMIQVTGFLDNEYFYVDFRDY 600  
QY 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGSKTGVSIQVAVMLKEKADSCKEKALM 660  
Db 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGSKTGVSIQVAVMLKEKADSCKEKALM 660  
QY 661 SELKQMTHLGHNDINVLIGACTLSGPVYLIPEYCCYGDLLNLYRSKREKPHRTWTIFK 720  
Db 661 SELKQMTHLGHNDINVLIGACTLSGPVYLIPEYCCYGDLLNLYRSKREKPHRTWTIFK 720  
QY 721 EHNFSYPTFOAHSNNSMFGSREVQLHPDLQSLSGFNNGSIHSEDEIEYENOKELABEE 780  
Db 721 EHNFSYPTFOAHSNNSMFGSREVQLHPDLQSLSGFNNGSIHSEDEIEYENOKELABEE 780  
QY 781 EDLNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGKVVKI CDFGLARDIL 840  
Db 781 EDLNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGKVVKI CDFGLARDIL 840  
QY 841 SPSSYVVRGNARLPVKWMAPESLFEGITYTKSDVNSYGILLWEIFSLGVNPPYGPV 900  
Db 841 SPSSYVVRGNARLPVKWMAPESLFEGITYTKSDVNSYGILLWEIFSLGVNPPYGPV 900  
QY 901 FYKLIQSGFKMEQPFYATGEGYFVMSQWAFDSRKPSFPNLTSLFLGCOLAEAEACIRT 960  
Db 901 FYKLIQSGFKMEQPFYATGEGYFVMSQWAFDSRKPSFPNLTSLFLGCOLAEAEACIRT 960  
QY 961 SIHLPKQAAPOQRGGLRAQSPQVKIHRERS 992  
Db 961 SIHLPKQAAPOQRGGLRAQSPQVKIHRERS 992

RESULT 3  
ID AAR67535 standard; Protein; 992 AA.  
XX AAR67535;  
XX AC  
XX DT 25-MAR-2003 (updated)  
XX DT 04-JUL-1995 (first entry)  
XX DE  
XX DE Mouse flk-2.  
XX KW Fetal liver kinase-2; flk-2; protein tyrosine-kinase receptor;  
XX KW hematopoiesis; stem cell.  
XX OS Mus sp.  
XX FH Key  
XX FT Peptide  
XX FT Domain

FT Region /label= Extracellular\_receptor\_domain  
FT 545..564  
FT /label= Transmembrane\_region  
FT Domain 565..992  
FT /label= Intracellular\_catalytic\_domain  
XX  
PN US367057-A.  
XX  
PD 22-NOV-1994.  
XX  
PF 30-APR-1993; 93US-0055269.  
XX  
PR 02-APR-1991; 91US-0679666.  
PR 28-JUN-1991; 91US-0728913.  
PR 15-NOV-1991; 91US-0793065.  
PR 24-DEC-1991; 91US-0813593.  
PR 26-JUN-1992; 92US-0906397.  
PR 12-NOV-1992; 92US-0975049.  
PR 19-NOV-1992; 92US-0977451.  
PR 30-APR-1993; 93US-0055269.  
XX  
PA (UYPR-) UNIV PRINCETON.  
XX  
PI Lemischka IR;  
XX  
DR WPI; 1995-005894/01.  
DR N-PSDB; AAQ79068.  
XX  
PT Murine flk-2 receptor protein tyrosine kinase - used to stimulate  
PT proliferation and/or stimulation of primitive mammalian  
PT haematopoietic stem cells in vitro or in vivo.  
XX  
XX Claim 1; Fig. 1A-1F; 69pp; English.  
XX  
XX cDNAs encoding receptor protein tyrosine-kinases, mouse fetal liver  
CC kinase-2 (flk-2), human flk-2 and mouse flk-1 are given in AAQ79068-70,  
CC respectively, and the deduced amino acid sequences in AAQ67535-37,  
CC respectively.  
CC (Updated on 25-MAR-2003 to correct PF field.)  
XX  
SQ Sequence 992 AA;  
Query Match 100.0%; Score 5264; DB 16; Length 992;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRALQRSRRLLLVLSWMILETVTNQDLPVVKVLSHENNGSSAGKPSRYMRVRS 60  
DB 1 MRALQRSRRLLLVLSWMILETVTNQDLPVVKVLSHENNGSSAGKPSRYMRVRS 60  
QY 61 PEDLOCTPRRQSGTVEAATVEVAESGSITLQVQLATPGDLSCLVFKHSSILGCCPHFD 120  
DB 61 PEDLOCTPRRQSGTVEAATVEVAESGSITLQVQLATPGDLSCLVFKHSSILGCCPHFD 120  
QY 121 LQNRGIVSMALNVTTQAGEYLLHTQSERANVTFLTVNVRDTQYLVLRPVRKQENQ 180  
DB 121 LQNRGIVSMALNVTTQAGEYLLHTQSERANVTFLTVNVRDTQYLVLRPVRKQENQ 180  
QY 181 DALLCISEGVPEPTVEWVLCSSHRESCKEGRVVRKEEVLHELFGTDIRCCARNALGR 240  
DB 181 DALLCISEGVPEPTVEWVLCSSHRESCKEGRVVRKEEVLHELFGTDIRCCARNALGR 240  
QY 241 ECTKLFITDINOAPQSTLPOLFLKVGEPMLWIRCKAIHVNHGFLTWELEDKALEEGSYFE 300  
DB 241 ECTKLFITDINOAPQSTLPOLFLKVGEPMLWIRCKAIHVNHGFLTWELEDKALEEGSYFE 300  
QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYYTCCSSSKHPSQSALVTILEKGFINATSSQEEY 360  
DB 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYYTCCSSSKHPSQSALVTILEKGFINATSSQEEY 360  
QY 361 EIDPYEKFCFVRFAKYPRICTWIFSQASFPCQORGLDGYISIKFCDHKNKPGYIFY 420  
DB 361 EIDPYEKFCFVRFAKYPRICTWIFSQASFPCQORGLDGYISIKFCDHKNKPGYIFY 420

RESULT 4

AAQ97418  
ID AAR97418 standard; Protein; 992 AA.  
XX  
AC AAR97418;  
XX  
XX 25-MAR-2003 (updated)  
DT 11-DEC-1996 (first entry)  
XX  
DE Human foetal liver kinase 2.  
XX  
KW Human; foetal liver kinase 2; flk-2; protein tyrosine kinase;  
KW monoclonal; antibody; extracellular domain; receptor assay;  
KW haematopoietic stem cell; ligand; stimulation; proliferation;  
KW differentiation; treatment; anaemia; bone marrow damage;  
KW cancer chemotherapy; radiation.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..27 /label= sig\_peptide  
FT Peptide 28..992 /label= mat\_peptide  
FT Domain 28..543 /label= extracellular\_domain  
FT Domain 544..563 /label= transmembrane\_domain  
FT Domain 564..992 /label= intracellular\_domain  
XX

QY 421 AENDDAQFTKFTLNIRKKPQVLANASASQSSDGYPLPSWTWKCKSDKSPNCTEETIP 480  
DB 421 AENDDAQFTKFTLNIRKKPQVLANASASQSSDGYPLPSWTWKCKSDKSPNCTEETIP 480  
QY 481 EGVNKKANRKVFQGVSSSTLNMEAGKGLAVKCCAYNSMGTSCETIFLNSPGPPFFIQ 540  
DB 481 EGVNKKANRKVFQGVSSSTLNMEAGKGLAVKCCAYNSMGTSCETIFLNSPGPPFFIQ 540  
QY 541 DNISFYATIGLCLPFIVVLIVLICHYKQKQFYESQLOMIQVTGPLDNIEFYVDFRDY 600  
DB 541 DNISFYATIGLCLPFIVVLIVLICHYKQKQFYESQLOMIQVTGPLDNIEFYVDFRDY 600  
QY 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISTGWSIQVAVKMLKEKADSCKEALM 660  
DB 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISTGWSIQVAVKMLKEKADSCKEALM 660  
QY 661 SELKMWTHLGHNDINVLIGACTLGGPVYLIIEYCCYGDLLNLYRSKRKFHRTWTET 720  
DB 661 SELKMWTHLGHNDINVLIGACTLGGPVYLIIEYCCYGDLLNLYRSKRKFHRTWTET 720  
QY 721 EHNFSYPTFOAHSNSMPSGSRVQLHPPLDQLSGFNNGSIHSEDEIEYENQKRLAEEEE 780  
DB 721 EHNFSYPTFOAHSNSMPSGSRVQLHPPLDQLSGFNNGSIHSEDEIEYENQKRLAEEEE 780  
QY 781 EDLNVLTTFEDLLCFAYQVAKGMEFLEFKSCVHRDLAARNVLTTHGKVKVVICDFGLARDIL 840  
DB 781 EDLNVLTTFEDLLCFAYQVAKGMEFLEFKSCVHRDLAARNVLTTHGKVKVVICDFGLARDIL 840  
QY 841 SDSSYVVRGNARLPVKWMAPELSLFGIYTIKSDVMSYGLLMEIFSLGVNPNYPGIPVDAN 900  
DB 841 SDSSYVVRGNARLPVKWMAPELSLFGIYTIKSDVMSYGLLMEIFSLGVNPNYPGIPVDAN 900  
QY 901 FYKLQSGFKMEQPPVATEGIVFMQSCWAFDSRKRPSFPNLTSLGCOLAAEEACIRT 960  
DB 901 FYKLQSGFKMEQPPVATEGIVFMQSCWAFDSRKRPSFPNLTSLGCOLAAEEACIRT 960  
QY 961 SIHLPKQAAPQORGGRLAQSQRQVKIHRERS 992  
DB 961 SIHLPKQAAPQORGGRLAQSQRQVKIHRERS 992

PN US548065-A.  
 XX 20-AUG-1996.  
 XX 31-OCT-1994; 94US-0252517.  
 XX 19-NOV-1992; 92US-0977451.  
 PR 02-APR-1991; 91US-0679666.  
 PR 28-JUN-1991; 91US-0728913.  
 PR 15-NOV-1991; 91US-0793065.  
 PR 24-DEC-1991; 91US-0813593.  
 PR 26-JUN-1992; 92US-0906397.  
 PR 12-NOV-1992; 92US-0975049.  
 PR 30-APR-1993; 93US-0055269.  
 PR 31-OCT-1994; 94US-0252517.  
 XX (UYPR-) UNIV PRINCETON.  
 PA Lemischka IR;  
 XX WPI; 1996-392678/39.  
 DR N-PSDB; AAT38733.  
 XX Anti-fetal liver kinase 2 (flk-2) antibodies - useful in assays,  
 PT for isolating haematopoietic stem cells expressing receptor and for  
 PT obtaining ligands  
 XX Claim 1; Columns 27-34; 50pp; English.  
 XX The present sequence is human foetal liver kinase 2 (flk-2), a  
 CC protein tyrosine kinase. Isolated antibodies, pref. monoclonal,  
 CC raised against the extracellular portion of flk-2 can be used to  
 CC assay for flk receptors on the surface of primitive haematopoietic  
 CC stem cells, and to isolate positive cells. The antibodies can also  
 CC be used as, or to obtain ligands, which stimulate the proliferation  
 CC and/or differentiation of stem cells. The ligands can be used, e.g.  
 CC for treating anaemia, or bone marrow damage resulting from cancer  
 CC chemotherapy, or radiation.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX Sequence 992 AA;  
 SQ  
 Query Match 100.0%; Score 5264; DB 17; Length 992;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MEALQSRDRRLLLLVLSVMILETVTNQDLPVVKCVLISHENNGSSAGKPSYRMVGRS 60  
 DB 1 MEALQSRDRRLLLLVLSVMILETVTNQDLPVVKCVLISHENNGSSAGKPSYRMVGRS 60  
 QY 61 PEDLOCTPRQSEGTVEAATVEVAESGSIITQVQLATPGDLSCLVFKHSSLGCOQPHD 120  
 DB 61 PEDLOCTPRQSEGTVEAATVEVAESGSIITQVQLATPGDLSCLVFKHSSLGCOQPHD 120  
 QY 121 LQNRGIVSMALNVTTQAGEYLLHIQSERANYTVLFTVNVDRDTQLYVLRPPYRKMENQ 180  
 DB 121 LQNRGIVSMALNVTTQAGEYLLHIQSERANYTVLFTVNVDRDTQLYVLRPPYRKMENQ 180  
 QY 181 DALLCISGVPEPTVEVWVLCSSHRESCKEKGPAVVRKEKVELHFGTDIRCCARNALGR 240  
 DB 181 DALLCISGVPEPTVEVWVLCSSHRESCKEKGPAVVRKEKVELHFGTDIRCCARNALGR 240  
 QY 241 ECTKLTIDLNQAPOSTLPQLFKVGEPLWIRKCAIHVNVHGFGLTWELDKALEGSYPE 300  
 DB 241 ECTKLTIDLNQAPOSTLPQLFKVGEPLWIRKCAIHVNVHGFGLTWELDKALEGSYPE 300  
 QY 301 MSTYSTNRTMIRILLAFVSSVGRNDGYTCSSSKHPSQSALVTILEKGFINATSSQEEY 360  
 DB 301 MSTYSTNRTMIRILLAFVSSVGRNDGYTCSSSKHPSQSALVTILEKGFINATSSQEEY 360  
 QY 361 EIDPYEKCFCFSVRFRKAYPRIRCTWIFPSQAFPCBQKGLDGYISIKFCDHKNKPGEIYF 420  
 DB 361 EIDPYEKCFCFSVRFRKAYPRIRCTWIFPSQAFPCBQKGLDGYISIKFCDHKNKPGEIYF 420

QY 421 AENDDAQFTKMTLNIRKKPQVLANASASQSSDGYPLPSWTWKCKSDKSPNCTBEIP 480  
 DB 421 AENDDAQFTKMTLNIRKKPQVLANASASQSSDGYPLPSWTWKCKSDKSPNCTBEIP 480  
 QY 481 EGVNKKANRKFVGQWVSSSTLNMSBAGKLLVKKCCAYNSMGTSCTIFLNSPGPPFFIQ 540  
 DB 481 EGVNKKANRKFVGQWVSSSTLNMSBAGKLLVKKCCAYNSMGTSCTIFLNSPGPPFFIQ 540  
 QY 541 DNISFYATIGLCPLPFIWLVILVILCHYKQKQFRYESQLOMIQVGTGPDNEFYVDFRDEY 600  
 DB 541 DNISFYATIGLCPLPFIWLVILVILCHYKQKQFRYESQLOMIQVGTGPDNEFYVDFRDEY 600  
 QY 601 DLKWEPPRENLEFGKVLGSGAFGRVNNATAYGTSKTVSIQAVKMLKEKADSCKEALM 660  
 DB 601 DLKWEPPRENLEFGKVLGSGAFGRVNNATAYGTSKTVSIQAVKMLKEKADSCKEALM 660  
 QY 661 SELKMMTHLGHDNIIVNLLGACTLSPVVLIFEYCCYGDLLNLYLSKREKPHRTWTIFK 720  
 DB 661 SELKMMTHLGHDNIIVNLLGACTLSPVVLIFEYCCYGDLLNLYLSKREKPHRTWTIFK 720  
 QY 721 EHNFSYPTFOAHSNMSMPGSRREVQLHPPDLQSLGFNGNSIHSEDEIEYENQKRLAEEEE 780  
 DB 721 EHNFSYPTFOAHSNMSMPGSRREVQLHPPDLQSLGFNGNSIHSEDEIEYENQKRLAEEEE 780  
 QY 781 EDNLVLTPEDLICFAYQVAKGMFLEFKSCVHRDLAARNVLTHTGKVKVVICDFGLARDIL 840  
 DB 781 EDNLVLTPEDLICFAYQVAKGMFLEFKSCVHRDLAARNVLTHTGKVKVVICDFGLARDIL 840  
 QY 841 SDSSYVVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNPPYGPVVDAN 900  
 DB 841 SDSSYVVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNPPYGPVVDAN 900  
 QY 901 FYKLQSGFKMEQPPFYATGIIYFVMSQWAFDSRKEPSPNLTSLFLGCQLAAEEACIRT 960  
 DB 901 FYKLQSGFKMEQPPFYATGIIYFVMSQWAFDSRKEPSPNLTSLFLGCQLAAEEACIRT 960  
 QY 961 SIHLPKQAAPQORGGGLRAOSPORQVKIHRERS 992  
 DB 961 SIHLPKQAAPQORGGGLRAOSPORQVKIHRERS 992  
 RESULT 5  
 AAW19874  
 ID AAW19874 standard; Protein; 992 AA.  
 XX AAW19874;  
 AC AC  
 DT 25-MAR-2003 (updated)  
 DT 19-AUG-1997 (first entry)  
 DE Murine flk-2 receptor.  
 XX Human; fetal liver kinase 2; flk2; receptor protein tyrosine kinase;  
 KW ptk; liver; spleen; thymus; adult; brain; bone marrow; primitive portion;  
 KW haematopoietic hierarchy; extracellular domain; soluble form; ligand;  
 KW proliferation; differentiation; mammalian; haematopoietic stem cell;  
 KW macrocytic anaemia; aplastic anaemia; cancer; radiation; mouse.  
 XX Mus musculus.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..27  
 FT /note= "Signal peptide"  
 FT Domain 28..544  
 FT /note= "Extracellular receptor domain"  
 FT Region 545..564  
 FT /note= "transmembrane region"  
 FT Domain 565..992  
 FT /note= "Intracellular catalytic domain"  
 FT Active-site 618..623  
 FT /note= "Catalytic sub-domain"  
 FT Active-site 811..819



FT Active-site /note= "Catalytic sub-domain"  
 FT 832..834  
 FT Active-site /note= "Catalytic sub-domain"  
 FT 857..862  
 FT Active-site /note= "Catalytic sub-domain"  
 FT 872..878  
 FT Active-site /note= "Catalytic sub-domain"  
 FT 709..785  
 FT Peptide /note= "signature sequence characteristic of flk-2"  
 FT 709..785  
 XX US5621090-A.  
 XX 15-APR-1997.  
 XX 26-JUN-1992; 92US-0906397.  
 XX 26-JUN-1992; 92US-0906397.  
 PR 02-APR-1991; 91US-0679666.  
 PR 28-JUN-1991; 91US-0728913.  
 PR 15-NOV-1991; 91US-0793065.  
 PR 24-DEC-1991; 91US-0813593.  
 XX (UYPR-) UNIV PRINCETON.  
 XX Lemischka IR;  
 XX WPI; 1997-235228/21.  
 DR N-PSDB; AAT72118.  
 XX Protein containing the extracellular domain of human flk-2 - used  
 PT for identification of primitive haematopoietic cell proliferation  
 PT and differentiation stimulatory ligands, e.g. for treating anaemia  
 XX Disclosure; Fig 1A; 55pp; English.  
 PS This sequence represents murine fetal liver kinase 2 (flk2). flk-2 is  
 CC a receptor protein tyrosine kinase (PTK) and is important in transducing  
 CC putative self-renewal signals from the environment. flk-2 is expressed  
 CC in fetal liver, spleen, fetal thymus, adult brain and adult bone marrow,  
 CC and it is thought that flk-2 is expressed in the entire primitive portion  
 CC of the haematopoietic hierarchy. The invention concerns a recombinant  
 CC nucleic acid, preferably mRNA, which encodes a protein containing only  
 CC the extracellular domain of human flk-2 and lacking the flk-2 intra-  
 CC cellular catalytic domain. The resultant protein represents a soluble  
 CC form of flk-2 which is used to isolate specific ligands for flk-2. These  
 CC ligands can be used to stimulate proliferation and/or differentiation of  
 CC mammalian haematopoietic stem cells, in vivo or in vitro, e.g. for  
 CC treatment of macrocytic or aplastic anaemia or bone marrow damage caused  
 CC by cancer treatment or radiation.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX Sequence 992 AA;  
 SQ

Query Match 100.0%; Score 5264; DB 18; Length 992;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALQSRDRRLLLVLSVMIETVNDLPVVKVLSHNNSSAGKPSRYMRVRS 60  
 DB 1 MRALQSRDRRLLLVLSVMIETVNDLPVVKVLSHNNSSAGKPSRYMRVRS 60  
 QY 61 PEDLQTPRQSGTGYEATVEAESGSIITLQVQLATPGDLSCLWFKHSSLCQPHFD 120  
 DB 61 PEDLQTPRQSGTGYEATVEAESGSIITLQVQLATPGDLSCLWFKHSSLCQPHFD 120  
 QY 121 LQNRGIVSMAILNVTTQAGEYLLHIQSERANYTLFTVNVRTQLYVLRPFYRWENQ 180  
 DB 121 LQNRGIVSMAILNVTTQAGEYLLHIQSERANYTLFTVNVRTQLYVLRPFYRWENQ 180  
 QY 181 DALLCISEGVPEPTVWVLCSSHRESCKEGPAVRKEEVLTGTDIRCCARNALGR 240  
 DB 181 DALLCISEGVPEPTVWVLCSSHRESCKEGPAVRKEEVLTGTDIRCCARNALGR 240

QY 241 ECTKLTFTIDLNQAPQSTLPOLFLKVGEPILMIRKCAIHVNHGFGILTWELEKALBEGSYFE 300  
 DB 241 ECTKLTFTIDLNQAPQSTLPOLFLKVGEPILMIRKCAIHVNHGFGILTWELEKALBEGSYFE 300  
 QY 301 MSTYSTNRTMIRILLAFVSSVGRNDGYTCSKSKHPSQSALVTLILKGFNATSSOEY 360  
 DB 301 MSTYSTNRTMIRILLAFVSSVGRNDGYTCSKSKHPSQSALVTLILKGFNATSSOEY 360  
 QY 361 EIDPYEKFCFSVRKAYPRIRCTWIFSQASFPCEORGLDGYISIKFCDHKNKFGYIFY 420  
 DB 361 EIDPYEKFCFSVRKAYPRIRCTWIFSQASFPCEORGLDGYISIKFCDHKNKFGYIFY 420  
 QY 421 AENDDAQFTKMTINIRKPKOVLANASASQASCSDDGYPLPSWTWKCSKSPNCTBEIP 480  
 DB 421 AENDDAQFTKMTINIRKPKOVLANASASQASCSDDGYPLPSWTWKCSKSPNCTBEIP 480  
 QY 481 EGVNKKANRKFQGVQVSSSTLNMSBAGKLLVKCCAYNSMGTSCTIFLNSPGPPPIQ 540  
 DB 481 EGVNKKANRKFQGVQVSSSTLNMSBAGKLLVKCCAYNSMGTSCTIFLNSPGPPPIQ 540  
 QY 541 DNISFYATIGLCPLFFIVLVILVILCHYKQFRYESQLQMTQVTGPDNBYFYVDFRDY 600  
 DB 541 DNISFYATIGLCPLFFIVLVILVILCHYKQFRYESQLQMTQVTGPDNBYFYVDFRDY 600  
 QY 601 DLKWEPPRENLEFGKVLGSGAFGRVNNATAYGISTGVSIOQAVKMLKEKADSCKEKALM 660  
 DB 601 DLKWEPPRENLEFGKVLGSGAFGRVNNATAYGISTGVSIOQAVKMLKEKADSCKEKALM 660  
 QY 661 SELKMTLHGHDNIYNLIGACTLSGPVYLIFEVCCYGDLLNLYLSKREKPHRTWTIFK 720  
 DB 661 SELKMTLHGHDNIYNLIGACTLSGPVYLIFEVCCYGDLLNLYLSKREKPHRTWTIFK 720  
 QY 721 EHNFSYPTFOAHSNMPGSRREVQLHPDLQSLGFGNSIHSEDEIYEYENQKRLAESEE 780  
 DB 721 EHNFSYPTFOAHSNMPGSRREVQLHPDLQSLGFGNSIHSEDEIYEYENQKRLAESEE 780  
 QY 781 EDNLVTFEDLFCFAQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKIKCDFGLARDIL 840  
 DB 781 EDNLVTFEDLFCFAQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKIKCDFGLARDIL 840  
 QY 841 SDSSYVVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIFSLGVNYPYGPV 900  
 DB 841 SDSSYVVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIFSLGVNYPYGPV 900  
 QY 901 FYKLIQSGFKMEOPFYATGIFVMSQWAFDSRKPSPNLTSLFLGCOLAEAEACIRT 960  
 DB 901 FYKLIQSGFKMEOPFYATGIFVMSQWAFDSRKPSPNLTSLFLGCOLAEAEACIRT 960  
 QY 961 SIHLPKQAAPQQRGGLRAQSPQRQVKIHRERS 992  
 DB 961 SIHLPKQAAPQQRGGLRAQSPQRQVKIHRERS 992

RESULT 6  
 AAY08616  
 ID AAY08616 standard; Protein; 992 AA.  
 XX AC AAY08616;  
 XX DT 05-AUG-1999 (first entry)  
 XX Murine flk-2 protein.  
 XX Murine; flk-2; flk-1; cell isolation; fetal liver kinase; receptor;  
 KW monoclonal; polyclonal; antibody; tyrosine kinase.  
 XX Mus sp.  
 XX US5912133-A.  
 XX 15-JUN-1999.  
 XX 10-FEB-1998; 98US-0021324.

XX 19-NOV-1992; 92US-0977451.  
 PR 02-APR-1991; 91US-0679666.  
 PR 28-JUN-1991; 91US-0728913.  
 PR 15-NOV-1991; 91US-0793065.  
 PR 24-DEC-1991; 91US-0813593.  
 PR 26-JUN-1992; 92US-0906397.  
 PR 12-NOV-1992; 92US-0975049.  
 PR 30-APR-1993; 93US-0055269.  
 PR 31-OCT-1994; 94US-0252498.  
 PR 15-FEB-1996; 96US-0601891.  
 XX (UYPR-) UNIV PRINCETON.  
 XX Lemischka IR;  
 XX WPI; 1999-357194/30.  
 DR N-PSDB; AAX77514.  
 XX  
 PT Isolating hematopoietic cells expressing fetal liver kinase 1  
 PT receptors  
 XX  
 PS Disclosure; Fig 1a; 59pp; English.  
 XX  
 CC This invention describes a novel method of isolating cells expressing  
 CC fetal liver kinase 1 (flk-1) receptors on their surface and comprises  
 CC binding the cells to a polyclonal or monoclonal antibody specific to  
 CC the flk-1 receptor and isolating the cells that have bound to the  
 CC antibody. The method can be used to isolate hematopoietic stem cells in  
 CC any mammal but preferably a rat, mouse, rabbit or human. The proteins of  
 CC the invention belong to the receptor protein family. This sequence  
 CC represents the murine flk-2 protein which is used in the method of  
 CC the invention.  
 XX  
 XX Sequence 992 AA:  
 SQ  
 Query Match 100.0%; Score 5264; DB 20; Length 992;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MRALAQRDRLLLLVLSVMILETVTNQDLPVVKCVLISHENNGSSAGKPSRYRMVRGS 60  
 DB 1 MRALAQRDRLLLLVLSVMILETVTNQDLPVVKCVLISHENNGSSAGKPSRYRMVRGS 60  
 QY 61 PEDLOCTPRROSEGVYEAATVEAEGSITLQVOLATPGDLSCLVWFKHSLSGCQPHD 120  
 DB 61 PEDLOCTPRROSEGVYEAATVEAEGSITLQVOLATPGDLSCLVWFKHSLSGCQPHD 120  
 QY 121 LONRGIVSMALNVTTQAGEYLLHIQSERANYTVLFTVNRDTPGLSCLWVFKHSLSGCQPHD 180  
 DB 121 LONRGIVSMALNVTTQAGEYLLHIQSERANYTVLFTVNRDTPGLSCLWVFKHSLSGCQPHD 180  
 QY 181 DALLCISRGVPEPTVEWVLCSSHRESCKEKGPAVVRKEKVLHFGTDIRCCARNALGR 240  
 DB 181 DALLCISRGVPEPTVEWVLCSSHRESCKEKGPAVVRKEKVLHFGTDIRCCARNALGR 240  
 QY 241 ECTKLFTIDLNAQPOSTLPQLFKVGEPLWIRCKAIHVNHGFGLTWELEDKALEEGSYFE 300  
 DB 241 ECTKLFTIDLNAQPOSTLPQLFKVGEPLWIRCKAIHVNHGFGLTWELEDKALEEGSYFE 300  
 QY 301 MSTYSTNRMTIRILLAFVSSVGRNDTGYTSSSKHPSQSALVTILEKGFNATSSQBEY 360  
 DB 301 MSTYSTNRMTIRILLAFVSSVGRNDTGYTSSSKHPSQSALVTILEKGFNATSSQBEY 360  
 QY 361 BIDPYEFCFVRFKAYPRICTWIFSOASPCGEORGLDGYISKFCDHKNKPGYEIFY 420  
 DB 361 BIDPYEFCFVRFKAYPRICTWIFSOASPCGEORGLDGYISKFCDHKNKPGYEIFY 420  
 QY 421 AENDDAQFTKMTLNIRKKPQVLANASASQASCSDDGYPLPSWTWKKCSKSPNCTEIP 480  
 DB 421 AENDDAQFTKMTLNIRKKPQVLANASASQASCSDDGYPLPSWTWKKCSKSPNCTEIP 480  
 QY 481 EGVNKKANRKVFQGVSSSTLNMSBAGKLLVKCCAYNSMGTSCTETIFLNSPGPFPIQ 540

DB 481 EGVNKKANRKVFQGVSSSTLNMSBAGKLLVKCCAYNSMGTSCTETIFLNSPGPFPIQ 540  
 QY 541 DNISFYATIGCLPFIIVVLIVLICHYKQPRYESQLOMIQVTGPLDNEYFYVDFRDY 600  
 DB 541 DNISFYATIGCLPFIIVVLIVLICHYKQPRYESQLOMIQVTGPLDNEYFYVDFRDY 600  
 QY 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISTGYSIQVAVOMLKEKADSCKEALM 660  
 DB 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISTGYSIQVAVOMLKEKADSCKEALM 660  
 QY 661 SELKXMTLGHHDNIVNLGACTLSGPVYLIFEYCCYGDLLNLYRSKREKPHRTWTEIFK 720  
 DB 661 SELKXMTLGHHDNIVNLGACTLSGPVYLIFEYCCYGDLLNLYRSKREKPHRTWTEIFK 720  
 QY 721 EHNFSSTYPTFOAHNSNMPGSRREVQLHPPLDQLSGFNNGSIHSEDEIEYENQKRLAESEE 780  
 DB 721 EHNFSSTYPTFOAHNSNMPGSRREVQLHPPLDQLSGFNNGSIHSEDEIEYENQKRLAESEE 780  
 QY 781 EDLNVLPEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKVLCDFGLARDIL 840  
 DB 781 EDLNVLPEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKVLCDFGLARDIL 840  
 QY 841 SDSSVVRGNARLPVKWMAPELPEGIYTIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900  
 DB 841 SDSSVVRGNARLPVKWMAPELPEGIYTIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900  
 QY 901 FYKLITQSGFKNEQPFYATGEGYFVWQSCWAFDSRKRPSFPNLTSLFGQLAEAEACIRT 960  
 DB 901 FYKLITQSGFKNEQPFYATGEGYFVWQSCWAFDSRKRPSFPNLTSLFGQLAEAEACIRT 960  
 QY 961 SIHLPKQAAPQORGLRAQSPQRQVKIHRERS 992  
 DB 961 SIHLPKQAAPQORGLRAQSPQRQVKIHRERS 992  
 RESULT 7  
 ABG70915  
 ID ABG70915 standard; Protein; 992 AA.  
 XX AC ABG70915;  
 XX DT 09-DEC-2002 (first entry)  
 XX Mouse receptor protein tyrosine kinase, FLK-2.  
 DE Mouse; FLK-2; foetal liver kinase; antianaemic; enzyme;  
 KW ophthalmological; receptor protein tyrosine kinase; aplastic anaemia;  
 KW primitive haematopoietic cell; stem cell; macrocytic anaemia;  
 KW bone marrow damage; cancer chemotherapy.  
 XX OS Mus sp.  
 XX Key Location/Qualifiers  
 FH Peptide 1..27  
 FT Protein /label= Signal\_peptide  
 FT Protein 28..992  
 XX /label= Mature\_FLK\_2  
 PN US2002119545-A1.  
 XX 29-AUG-2002.  
 XX 01-JUN-2001; 2001US-08721136.  
 XX 19-NOV-1992; 92US-0977451.  
 PR 30-APR-1993; 93US-0055269.  
 PR 31-OCT-1994; 94US-0252498.  
 PR 15-FEB-1996; 96US-0601891.  
 PR 10-FEB-1998; 98US-0021324.  
 PR 10-DEC-1998; 98US-0208786.  
 PR 02-APR-1991; 91US-0679666.  
 PR 28-JUN-1991; 91US-0728913.



PR 19-NOV-1992; 92US-0977451.  
 PR 30-APR-1993; 93US-0055269.  
 PR 31-OCT-1994; 94US-0252498.  
 PR 15-FEB-1996; 96US-0601891.  
 PR 10-FEB-1998; 98US-0021324.  
 PR 10-DEC-1998; 98US-0208786.  
 PR 02-APR-1991; 91US-0679666.  
 PR 28-JUN-1991; 91US-0728913.  
 PR 15-NOV-1991; 91US-0793065.  
 PR 24-DEC-1991; 91US-0813593.  
 PA (LEWIS) LEMISCHKA I R.  
 XX  
 XX Lemischka IR;  
 DR WPI; 2002-607237/65.  
 DR N-PSDB; AAD42483.  
 XX  
 XX New protein tyrosine kinase expressed in primitive hematopoietic cells  
 PT (HC) and not expressed in mature HC, and ligands for the protein, for  
 PT stimulating proliferation of primitive hematopoietic stem cells -  
 XX  
 PS Claim 37; Page 19-21; 64pp; English.  
 XX  
 CC The present invention relates to receptor protein tyrosine kinase (pTK)  
 CC expressed in primitive hematopoietic cells and not expressed in mature  
 CC hematopoietic cells, polynucleotides encoding such proteins and ligands  
 CC for the protein. Ligands which bind to pTK of the invention are useful  
 CC for stimulating the proliferation and/or differentiation of primitive  
 CC mammalian hematopoietic stem cells. The receptor pTK molecules are  
 CC useful for stimulating the self renewal of the totipotent hematopoietic  
 CC stem cell and to stimulate the development of all cells of hematopoietic  
 CC system both in vitro and in vivo. The ligands for the receptors act as  
 CC hematopoietic growth factors. The ability of the ligands to stimulate  
 CC proliferation of stem cells both in vitro and in vivo has important  
 CC therapeutic applications such as treating humans whose primitive stem  
 CC cells do not sufficiently undergo self-renewal. It is also useful in  
 CC conditions that occur when defects in hematopoietic stem cells or their  
 CC related growth factors depress the number of white blood cells such as  
 CC macrocytic and aplastic anaemia or bone marrow damage resulting from  
 CC cancer chemotherapy and radiation. The present sequence is murine  
 CC receptor pTK, flk-2.  
 XX  
 SQ Sequence 992 AA;  
 Query Match 100.0%; Score 5264; DB 23; Length 992;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MRALQSRDRLLLLVLSVMILETVTNQDLPVKVLSHENNGSSAGKPSRYMVRGS 60  
 DB 1 MRALQSRDRLLLLVLSVMILETVTNQDLPVKVLSHENNGSSAGKPSRYMVRGS 60  
 QY 61 PEDLOTPRQSEGVYEAATVEAEGSSITLQVQLATPGDLSCILWFKHSGLGQPHFD 120  
 DB 61 PEDLOTPRQSEGVYEAATVEAEGSSITLQVQLATPGDLSCILWFKHSGLGQPHFD 120  
 QY 121 LQNGIVSMALNTVETQAGYLLHIQSERANYTVLFTVNVDRDQLYVLRPRYFRKMEQ 180  
 DB 121 LQNGIVSMALNTVETQAGYLLHIQSERANYTVLFTVNVDRDQLYVLRPRYFRKMEQ 180  
 QY 181 DALLCISEGVPEPTVEWVLCSSHRESCKECPAVVRKEEKVHLFGTDIRCCARNALGR 240  
 DB 181 DALLCISEGVPEPTVEWVLCSSHRESCKECPAVVRKEEKVHLFGTDIRCCARNALGR 240  
 QY 241 ECTKLFTIDLNQAPQSTLPQLFLKVGEPPLWIRCKAIHVNHGFGLTWELEDKALEGYSYFE 300  
 DB 241 ECTKLFTIDLNQAPQSTLPQLFLKVGEPPLWIRCKAIHVNHGFGLTWELEDKALEGYSYFE 300  
 QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCCSSSKHPSQSALVTILEKGFNATSSQEEY 360  
 DB 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCCSSSKHPSQSALVTILEKGFNATSSQEEY 360

QY 361 EIDPYEKFCFSVRFKAYPRIRCTWIFSOASFPCEORGLDGYISKFCDHKNKRGGEYIFY 420  
 DB 361 EIDPYEKFCFSVRFKAYPRIRCTWIFSOASFPCEORGLDGYISKFCDHKNKRGGEYIFY 420  
 QY 421 AENDDAQFTKMFNLIRKKPQVLNANASASQSSDGYPLPSWTWKCKSDKSPNCTREIP 480  
 DB 421 AENDDAQFTKMFNLIRKKPQVLNANASASQSSDGYPLPSWTWKCKSDKSPNCTREIP 480  
 QY 481 EGVNKKANKRVFGQWVSSSTLANSEBAGKLLVKCCAYNSMGTSCTIFLNSPGFPFPIQ 540  
 DB 481 EGVNKKANKRVFGQWVSSSTLANSEBAGKLLVKCCAYNSMGTSCTIFLNSPGFPFPIQ 540  
 QY 541 DNISFYATIGLCPLPFIWLVILCHKYKQFRYESQLOMIOVTGPDNEFYVDYFRDYEY 600  
 DB 541 DNISFYATIGLCPLPFIWLVILCHKYKQFRYESQLOMIOVTGPDNEFYVDYFRDYEY 600  
 QY 601 DLKWEPPRENLEFGKVLGSGAFGRVNNATAYGISKTGVSIOQAVKMLKEKADSCKEKALM 660  
 DB 601 DLKWEPPRENLEFGKVLGSGAFGRVNNATAYGISKTGVSIOQAVKMLKEKADSCKEKALM 660  
 QY 661 SELKMMTHLGHNDINVLGACTLSGPVYLIFEYCCYGDLLNLYLSKREKFRHTWTIEFK 720  
 DB 661 SELKMMTHLGHNDINVLGACTLSGPVYLIFEYCCYGDLLNLYLSKREKFRHTWTIEFK 720  
 QY 721 EHNFSYPTFOAHSNMSMPGSRVQLHPDLQSLGNGNSIHSEDEIEYENQKRLAESEE 780  
 DB 721 EHNFSYPTFOAHSNMSMPGSRVQLHPDLQSLGNGNSIHSEDEIEYENQKRLAESEE 780  
 QY 781 EDLNLVTFEDLLCFAYQVAKGMFELEFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840  
 DB 781 EDLNLVTFEDLLCFAYQVAKGMFELEFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840  
 QY 841 SDSSYVVRGNARLPVKWAPESLFGIYTIKSDVMSYGILLWEIFSLGVNYPYGPVPDAN 900  
 DB 841 SDSSYVVRGNARLPVKWAPESLFGIYTIKSDVMSYGILLWEIFSLGVNYPYGPVPDAN 900  
 QY 901 FYKLQSGFKMEQPPYATEGIVFMQSWAFPSRKRPSFPNLTSLGCOLAEAEACIRT 960  
 DB 901 FYKLQSGFKMEQPPYATEGIVFMQSWAFPSRKRPSFPNLTSLGCOLAEAEACIRT 960  
 QY 961 SIHLPKQAAPQORGLRAQSPQRQVKIHRERS 992  
 DB 961 SIHLPKQAAPQORGLRAQSPQRQVKIHRERS 992  
 RESULT 9  
 AAR44994  
 ID AAR44994 standard; Protein; 992 AA.  
 XX  
 AC AAR44994;  
 XX  
 DT 25-MAR-2003 (updated)  
 XX  
 DT 27-JUN-1994 (first entry)  
 XX  
 DE Murine flk-2 receptor protein tyrosine kinase.  
 XX  
 KW Receptor protein tyrosine kinase; pTK family; foetal liver kinase;  
 KW primitive; totipotent; haematopoietic cell; stem cell; proliferation;  
 KW mflk; stromal cell.  
 XX  
 OS Mus musculus.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..27  
 FT /label= signal\_peptide  
 FT /note= "hydrophobic leader"  
 FT Protein 28..992  
 FT /label= flk-2  
 FT /note= "mature protein"  
 FT Domain 28..544  
 FT /label= extracellular\_domain  
 FT Region 545..564  
 FT /label= transmembrane\_region

FT	Domain	565..992	/label= intracellular_catalytic_domain
FT	Domain	618..623	/note= "catalytic subdomain"
FT	Domain	811..819	/note= "catalytic subdomain"
FT	Domain	832..834	/note= "catalytic subdomain"
FT	Domain	857..862	/note= "catalytic subdomain"
FT	Domain	872..878	/note= "catalytic subdomain"
XX	US270458-A.		
PN	14-DEC-1993.		
XX	19-NOV-1992;	92US-0977451.	
XX	02-APR-1991;	91US-0679666.	
PR	28-JUN-1991;	91US-0728913.	
PR	15-NOV-1991;	91US-0793065.	
PR	24-DEC-1991;	91US-0813593.	
PR	26-JUN-1992;	92US-0906397.	
PR	12-NOV-1992;	92US-0975049.	
PR	19-NOV-1992;	92US-0977451.	
XX	(UYPR-) UNIV PRINCETON.		
PA	Lemischka IR;		
XX	WPI; 1993-405021/50.		
DR	N-PSDB; AAQ53502.		
XX	Isolated nucleic acid molecules of hematopoietic stem cell		
PT	receptor flk-2 - encoding mammalian receptor protein tyrosine		
PT	kinases expressed in primitive haematopoietic cells		
XX	Disclosure; Fig 1a; 60pp; English.		
PS	Nucleic acid sequences coding for murine flk-2 (AAR44994) and		
XX	specified subfragments (i.e. the pre-protein, the mature protein		
CC	and the extracellular domain) of it are claimed. The flk-2		
CC	polypeptide is a protein tyrosine kinase expressed only in		
CC	primitive haematopoietic cells. The cDNA can be used to		
CC	recombinantly produce flk-2 for stimulating self-renewal of		
CC	totipotent stem cells and development of all haematopoietic cells.		
CC	(Updated on 25-MAR-2003 to correct PF field.)		
XX	Sequence 992 AA;		
SQ	Query Match	99.8%; Score 5256; DB 14; Length 992;	
	Best Local Similarity	99.9%; Pred. No. 0;	
	Matches 991; Conservative	0; Mismatches 1; Indels 0; Gaps 0;	
QY	1	MRALQSRDRRLLLVLSWMILETVTNQDLPVKVCLISHENNGSSACKPSSYRMVRGS	60
Db	1	MRALQSRDRRLLLVLSWMILETVTNQDLPVKVCLISHENNGSSACKPSSYRMVRGS	60
QY	61	PEDLOCTPRQSEGTVEAATVEVAESGSIITLQVQLATPGDLSCLWFKHSLGCPHPD	120
Db	61	PEDLOCTPRQSEGTVEAATVEVAESGSIITLQVQLATPGDLSCLWFKHSLGCPHPD	120
QY	121	LQNRGIVSMALINVTQAGEVLLHTQSERANYTVLFTVNVDRDTQLYVLRPFRKVENQ	180
Db	121	LQNRGIVSMALINVTQAGEVLLHTQSERANYTVLFTVNVDRDTQLYVLRPFRKVENQ	180
QY	181	DALLCISEGVPEPTVEWVLCSSHRESCKEEGPAVVRKEKVLHELFGTDIRCARNALGR	240
Db	181	DALLCISEGVPEPTVEWVLCSSHRESCKEEGPAVVRKEKVLHELFGTDIRCARNALGR	240
QY	241	ECTKLTIDLNQAPQSTLPQLFLKVGEPFLWIRCKAIHVNHGFGLTWELEDKALBEGSYFE	300

Db	241	ECTKLTIDLNQAPQSTLPQLFLKVGEPFLWIRCKAIHVNHGFGLTWELEDKALBEGSYFE	300
QY	301	MSTYSTNRMTIRILLAFVSSVGRNDTGYYTSSSKHPSQSALVITILEKGFINATSSQSEY	360
Db	301	MSTYSTNRMTIRILLAFVSSVGRNDTGYYTSSSKHPSQSALVITILEKGFINATSSQSEY	360
QY	361	EIDPYEKFCFSVRPKAYPRICRTWIFSQASPPCQORGLDGYISKFCDHKNKPGSEYIFY	420
Db	361	EIDPYEKFCFSVRPKAYPRICRTWIFSQASPPCQORGLDGYISKFCDHKNKPGSEYIFY	420
QY	421	AENDDAQFTKMTLNIIRKKQVLANASASQASCSGSDGYPLPSWTWKKCSDKSPNCTEETP	480
Db	421	AENDDAQFTKMTLNIIRKKQVLANASASQASCSGSDGYPLPSWTWKKCSDKSPNCTEETP	480
QY	481	EGVWKKANRKYFGQWSSSTLNMSKAGLLVKCCAINSGTSCETIFLNSPGPFPTQ	540
Db	481	EGVWKKANRKYFGQWSSSTLNMSKAGLLVKCCAINSGTSCETIFLNSPGPFPTQ	540
QY	541	DNISFYATIGLCLPPIVVLIVLI CHYKQKQFYSOLOMIQVTPGLDNEYFYVDFRDY	600
Db	541	DNISFYATIGLCLPPIVVLIVLI CHYKQKQFYSOLOMIQVTPGLDNEYFYVDFRDY	600
QY	601	DLKWEFFPRENLEFGKVLGSGAFGRVMNATAYGISKTGYSIQVAVKMLKEKADSCKEALM	660
Db	601	DLKWEFFPRENLEFGKVLGSGAFGRVMNATAYGISKTGYSIQVAVKMLKEKADSCKEALM	660
QY	661	SELKXMTHLGHHDNI VNLGACTLSGPVYLIFEYCCYGDLLNLYLRKREKPHRTWTEIFK	720
Db	661	SELKXMTHLGHHDNI VNLGACTLSGPVYLIFEYCCYGDLLNLYLRKREKPHRTWTEIFK	720
QY	721	EHNFSYPTFOAHNSNMPGSRVQLHPDLQSGFNCSIHSEDEIEYENOKRLAESEE	780
Db	721	EHNFSYPTFOAHNSNMPGSRVQLHPDLQSGFNCSIHSEDEIEYENOKRLAESEE	780
QY	781	EDLNVLTPEDDLCPAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKVVCDFGLARDIL	840
Db	781	EDLNVLTPEDDLCPAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKVVCDFGLARDIL	840
QY	841	SDSSYVVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIFSLGVNPPYGPVPDAN	900
Db	841	SDSSYVVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIFSLGVNPPYGPVPDAN	900
QY	901	FYKLIQSGFKMEQPEYATGEGYFVWQSCWAFDSRKRPSFPNLTSLFGCOLAEAEACIRT	960
Db	901	FYKLIQSGFKMEQPEYATGEGYFVWQSCWAFDSRKRPSFPNLTSLFGCOLAEAEACIRT	960
QY	961	SIHLPKQAAPQORGLRAQSPORQVKIHRERS	992
Db	961	SIHLPKQAAPQORGLRAQSPORQVKIHRERS	992
RESULT 10			
AAR31375			
ID	AAR31375 standard; Protein; 992 AA.		
XX	AAR31375;		
DT	25-MAR-2003 (updated)		
DT	25-JUN-1993 (first entry)		
DE	Murine flk-2.		
XX	Murine; receptor; protein; tyrosine kinase; ptk; primitive; mammalian;		
KW	hematopoietic cell; pHC; mature; mHC; fetal; liver kinase 2; flk-2;		
KW	liver; spleen; thymus; adult; brain; marrow; thymocyte; subset;		
KW	multipotential; T-lymphoid; lineage.		
OS	Mus musculus.		
XX	Key	Location/Qualifiers	
FT	Peptide	1..27	
FT	Domain	28..544	
		/note= "Hydrophobic leader sequence"	

FT	Region	/note= "Extracellular receptor domain"	Db	121	LQNRGIVSMALNVTTQTQAGEYLLHIQSERANYTVLFTVNVVDTQLYVLRRPYFRKWNQ	180
FT	Domain	545..564	Qy	181	DALLCISEGVPEPTVWVLCSSHRESCKEKGPAVVRKEEKVLHELFGTDIRCCARNALGR	240
FT	Domain	565..992	Db	181	DALLCISEGVPEPTVWVLCSSHRESCKEKGPAVVRKEEKVLHELFGTDIRCCARNALGR	240
FT	Domain	/note= "Intracellular catalytic domain"	Qy	241	ECKLFTIDLNQAPQSTLPQLFKVGEPLWIRCKAIHVNHGFGLTWELDKALEEGSYFE	300
FT	Domain	/note= "Catalytic subdomain #1"	Db	241	ECKLFTIDLNQAPQSTLPQLFKVGEPLWIRCKAIHVNHGFGLTWELDKALEEGSYFE	300
FT	Domain	618..623	Qy	301	MSTYSTNRTMIRILLAPVSSVGRNDTGYTTCSSSKHPSQSALVTILEKGFINATSSQEEY	360
FT	Domain	811..819	Db	301	MSTYSTNRTMIRILLAPVSSVGRNDTGYTTCSSSKHPSQSALVTILEKGFINATSSQEEY	360
FT	Domain	/note= "Catalytic subdomain #3"	Qy	361	EIDPYEKFCFSVRPKAYPRIRCTWIFSOASPCGEORGLDGYISIKFCDHKNKPGYEIFY	420
FT	Domain	832..834	Db	361	EIDPYEKFCFSVRPKAYPRIRCTWIFSOASPCGEORGLDGYISIKFCDHKNKPGYEIFY	420
FT	Domain	/note= "Catalytic subdomain #4"	Qy	421	AENDDAQPTKFTLNIRKKPOVLANASASQASQSSDGYPLPSWTWKKCDKSPNCTEETIP	480
FT	Domain	857..862	Db	421	AENDDAQPTKFTLNIRKKPOVLANASASQASQSSDGYPLPSWTWKKCDKSPNCTEETIP	480
FT	Domain	/note= "Catalytic subdomain #5"	Qy	481	EGVWKKANKRVFGQWVSSSTLNMSKAGKLLVKCCAYNSMGTSCTETIFLNSPGPPFFIQ	540
FT	Domain	872..878	Db	481	EGVWKKANKRVFGQWVSSSTLNMSKAGKLLVKCCAYNSMGTSCTETIFLNSPGPPFFIQ	540
FT	Domain	/note= "Catalytic subdomain #6"	Qy	541	DNISFYATIGLCPLFIIVLVLICHYKKQPRYESQLOMIQVTPGLDNEYFYVDFRDYEY	600
FT	Region	709..785	Db	541	DNISFYATIGLCPLFIIVLVLICHYKKQPRYESQLOMIQVTPGLDNEYFYVDFRDYEY	600
FT	Region	/note= "Signature sequence characteristic of flk-2"	Qy	601	DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISTGVSIOVAVKMLKEKADSCKEALM	660
PN	W09300349-A1.		Db	601	DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISTGVSIOVAVKMLKEKADSCKEALM	660
PD	07-JAN-1993.		Qy	661	SELKMWTHLGHNDINVLGACTLSGPPVLIFFEYCCYGDLLNLYLSRKEKFKHRTWTEIFK	720
PD	26-JUN-1992;	92WO-US05401.	Db	661	SELKMWTHLGHNDINVLGACTLSGPPVLIFFEYCCYGDLLNLYLSRKEKFKHRTWTEIFK	720
PF	28-JUN-1991;	91US-0728913.	Qy	721	EHNPSYPTQAHNSNMPGSRREVQLHPPDLQSLGFGNGNSIHSDEIEYENQKRLAESEE	780
PF	15-NOV-1991;	91US-0793065.	Db	721	EHNPSYPTQAHNSNMPGSRREVQLHPPDLQSLGFGNGNSIHSDEIEYENQKRLAESEE	780
PR	24-DEC-1991;	91US-0813593.	Qy	781	EDLNVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL	840
PR	02-APR-1992;	92WO-US02750.	Db	781	EDLNVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL	840
XX	(UYPR-) UNIV PRINCETON.		Qy	841	SDSSYVVRGNARLPVKWMAPESEFEGIYTIKSDVMSYGILLWEIFSLGVNPGIPVDAN	900
XX	Lemischka IR;		Db	841	SDSSYVVRGNARLPVKWMAPESEFEGIYTIKSDVMSYGILLWEIFSLGVNPGIPVDAN	900
XX	WPI; 1993-036323/04.		Qy	901	FYKLIQSGFKWQEPFYVATEGIYFVMQSCWAFDSRKRPSFPNLTSLFGCQLAEAEACIRT	960
XX	N-PSDB; AAQ45247.		Db	901	FYKLIQSGFKWQEPFYVATEGIYFVMQSCWAFDSRKRPSFPNLTSLFGCQLAEAEACIRT	960
PT	Nucleic acid encoding receptor protein tyrosine kinase - allows		Qy	961	SIHLPKQAAPQQRGGLRAQSPQRQVKIHRERS	992
PT	development of ligands to stimulate proliferation and/or		Db	961	SIHLPKQAAPQQRGGLRAQSPQRQVKIHRERS	992
PT	differentiation of mammalian haematopoietic stem cells		Qy			
PS	Claim 5; Fig 1a; 78pp; English.		Db			
XX	This sequence represents a murine receptor protein tyrosine kinase		Qy			
XX	which belongs to a new functional class of protein tyrosine kinases		Db			
CC	(PTKs). PTKs in this class are expressed in primitive mammalian		Qy			
CC	hematopoietic (pHC) cells but not in mature hematopoietic cells (mHC).		Db			
CC	This protein is an example of a receptor pTK and is called fetal liver		Qy			
CC	kinase 2 (flk-2). flk-2 is expressed in fetal liver, spleen and		Db			
CC	thymus, and adult brain and marrow. Expression of flk-2 mRNA occurs		Qy			
CC	in the most primitive thymocyte subset, which is believed to be		Db			
CC	uncommitted. Therefore, thymocytes expressing flk-2 may be multi-		Qy			
CC	potential. flk-2 is the first receptor tyrosine kinase known to be		Db			
CC	expressed in the T-lymphoid lineage.		Qy			
CC	(Updated on 25-MAR-2003 to correct PN field.)		Db			
XX	Query Match	99.8%; Score 5256; DB 14; Length 992;	RESULT 11			
XX	Best Local Similarity	99.9%; Pred. No. 0;	AA28038			
XX	Matches 991; Conservative	0; Mismatches 1; Indels 0; Gaps 0;	ID	AA28038	standard; Protein; 992 AA.	
Qy	1 MRALAQRSDRLLLVLSWMILETVNQDLPIKCVLISHENNGSSAGKPSRYMRVRS	60	XX	AA28038;		
Db	1 MRALAQRSDRLLLVLSWMILETVNQDLPIKCVLISHENNGSSAGKPSRYMRVRS	60	XX	AC		
Qy	61 PEDLOCTPRQSGTVEATVVAESGSIITQVQLATPGDLCLWVFKHSSIGCCPHFD	120	DT	25-MAR-2003	(updated)	
Db	61 PEDLOCTPRQSGTVEATVVAESGSIITQVQLATPGDLCLWVFKHSSIGCCPHFD	120	DT	15-MAR-1993	(first entry)	
Qy	121 LQNRGIVSMALNVTTQTQAGEYLLHIQSERANYTVLFTVNVVDTQLYVLRRPYFRKWNQ	180	XX		Murine flk-2.	
XX			XX		Thymidine kinase; TK; haematopoietic; stem cells; proliferation;	
XX			KW		differentiation; progenitor cells; foetal liver kinase.	
XX			XX			

[illegible]



XX PN WO927062-A1.  
XX PD 12-OCT-1995.  
XX PF 23-MAR-1995; 95WO-US037118.  
XX PR 04-APR-1994; 94US-0222299.  
XX PA (GETH ) GENENTECH INC.  
XX PI Bennett BD, Broz SD, Matthews W, Zeigler FC;  
XX DR WPI; 1995-358636/46.  
XX DR N-PSDB; AAT00801.  
XX PT Agonist antibody which binds to flk2-flt3 tyrosine kinase receptor  
PT - enhances proliferation of haematopoietic stem cells, in the  
PT treatment of hypoplasia, anaemia, etc.  
XX PS Example 1; Page 38-40; 59pp; English.  
XX CC Murine stem cell tyrosine kinase receptor flk2 (also called flt3)  
CC (AAR1868) is encoded DNA (AAT00801) obtd. by RT-PCR amplification of  
CC RNA isolated from mid-gestation mouse foetal livers. An flk2/flt3  
CC extracellular domain-IgG1 Fc fusion protein was used to raise  
CC agonist antibodies able to bind to, and activate, flk2/flt3.  
CC Such antibodies can enhance repopulation of mature blood cell  
CC lineages following chemotherapy, radiotherapy, or bone marrow  
CC transplantation.  
XX SQ Sequence 1000 AA;  
Query Match 96.9%; Score 5102; DB 16; Length 1000;  
Best Local Similarity 96.7%; Pred. No. 0;  
Matches 968; Conservative 6; Mismatches 17; Indels 10; Gaps 2;  
1 MRALQSRDRLLLVLSVMILETNTQDLPIKCVLISHENNGSSAGKPSRYMVRGS 60  
1 MRALQSRDRLLLVLSVMILETNTQDLPIKCVLISHENNGSSAGKPSRYMVRGS 60  
61 PEDLOCTPRQSGTYVEAATVVAESGSIITLQVQLATPGDLSCLWVFKHSSLCGCPHF 120  
61 PEDLOCTPRQSGTYVEAATVVAESGSIITLQVQLATPGDLSCLWVFKHSSLCGCPHF 120  
121 LQNRGIVSMALLNVTTQAGEYLLHIQSEAAVTYVLTNNVDTQLYVLRPYFRKWNQ 180  
121 LQNRGIVSMALLNVTTQAGEYLLHIQSEAAVTYVLTNNVDTQLYVLRPYFRKWNQ 180  
181 DALLCISEGVPEPTVEWVLCSSHRECKEGPAVVRKEEVKVLHELFGTDIRCCARNALGR 240  
181 DALLCISEGVPEPTVEWVLCSSHRECKEGPAVVRKEEVKVLHELFGTDIRCCARNALGR 240  
241 ECTKLTIDLNOAQSTLPQLFKVGEPLWIRCKAIVNHGFLTWELDKALEEGSYFE 300  
241 ECTKLTIDLNOAQSTLPQLFKVGEPLWIRCKAIVNHGFLTWELDKALEEGSYFE 300  
301 MSTYSTNRTMIRILLAFVSSVGRNDTGYVTCSSSKHPSQSALVTILEKGFINATSSQEEY 360  
301 MSTYSTNRTMIRILLAFVSSVGRNDTGYVTCSSSKHPSQSALVTILEKGFINATSSQEEY 360  
361 EIDPYEKFCFSVRFKAYPIRCTWIFSOASFPCEQRLGEGDYSISKPCDHKNKPGYIFY 420  
361 EIDPYEKFCFSVRFKAYPIRCTWIFSOASFPCEQRLGEGDYSISKPCDHKNKPGYIFY 420  
421 AENDDAQFTKMTFLNIRKPKQVLANASASOACSSDGYPLPSMTWKCKSDKSPNCTEETIP 480  
421 AENDDAQFTKMTFLNIRKPKQVLANASASOACSSDGYPLPSMTWKCKSDKSPNCTEETIP 480  
481 EGVNKNKANRVFGQWSSSTLNMSEAGKLLVKKCAVNSMGTSCETIFLNSPGPPPTIQ 540  
481 EGVNKNKANRVFGQWSSSTLNMSEAGKLLVKKCAVNSMGTSCETIFLNSPGPPPTIQ 540

QY 541 DNISFYATIGLCCLPFIIVLVILVILCHYKQPRYBSQLOMQVTCPLDNEFYVDFRDY 600  
DB 541 DNISFYATIGLCCLPFIIVLVILVILCHYKQPRYBSQLOMQVTCPLDNEFYVDFRDY 600  
QY 601 DLKWEPPRENLEFGKVLGSGAFGRVMNATAYGISTGVSQVAVKMLKEKADSCKEALM 660  
DB 601 DLKWEPPRENLEFGKVLGSGAFGRVMNATAYGISTGVSQVAVKMLKEKADSCKEALM 660  
QY 661 SELKMMTHLGHNDINVLGACTLSGPGVYLIFECYCCYDGLLNLRLSKREKPHRTWTEIFK 720  
DB 661 SELKMMTHLGHNDINVLGACTLSGPGVYLIFECYCCYDGLLNLRLSKREKPHRTWTEIFK 720  
QY 721 EHNFSYPTFOAHSNNSMPCSGREVQLHPPDLQSLGFGNGSIHSEDEIEYENOKRLAE 780  
DB 721 EHNFSYPTFOAHSNNSMPCSGREVQLHPPDLQSLGFGNGSIHSEDEIEYENOKRLAE 780  
QY 781 EDNLVLTFFEDLLCFAYQVAKGMFLEPKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840  
DB 781 EDNLVLTFFEDLLCFAYQVAKGMFLEPKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840  
QY 841 SDSSYVVRGNARLPVKWMAPESEIFEGITTIKSDVMSYGILLWEIFSLGVNPPYGPVDAN 900  
DB 841 SDSSYVVRGNARLPVKWMAPESEIFEGITTIKSDVMSYGILLWEIFSLGVNPPYGPVDAN 900  
QY 901 FYKLIOGFKMEQPFYATGIIYFVMSQWAFDSKRPSPNLTSLGCOLAABEACIRT 960  
DB 901 FYKLIOGFKMEQPFYATGIIYFVMSQWAFDSKRPSPNLTSLGCOLAABEACIRT 960  
QY 961 -----SIHLPKQAAPOORGLRAQSPQOROVKIHRS 992  
DB 961 MGNVPEHPSIYQNRRLPSREAGS-EPPSPQAQVKIHRS 1000  
RESULT 13  
AAR75961  
ID AAR75961 standard; Protein; 993 AA.  
XX AAR75961;  
XX 25-MAR-2003 (updated)  
DT 29-DEC-1995 (first entry)  
XX Human STK-1.  
XX STK-1; receptor PTK; protein tyrosine kinase.  
XX Homo sapiens.  
FH Key Location/Qualifiers  
FT Peptide 1..23  
FT /label= signal  
FT /note= "unique PTK receptor sequence"  
FT Cleavage-site 25..26  
FT Domain 35..524  
FT /note= "Part of extracellular domain contg.  
FT 22 Cys residues marked on Fig. 1"  
FT Modified-site 43..45  
FT /label= potential N-linked glycosylation site  
FT Modified-site 100..102  
FT /label= see above  
FT Modified-site 151..153  
FT /label= see above  
FT Modified-site 250..252  
FT /label= see above  
FT Modified-site 306..308  
FT /label= see above  
FT Modified-site 323..325  
FT /label= see above  
FT Modified-site 351..353  
FT /label= see above  
FT Modified-site 354..356  
FT /label= see above  
FT Modified-site 473..475

FT Modified-site /label= see above  
 FT 502..504  
 FT /label= see above  
 FT 542..562  
 FT /label= transmembrane spanning region  
 FT 708..782  
 FT /label= kinase insert region  
 FT 617..622  
 FT /label= ATP binding domain  
 FT 835..840  
 FT /label= WMAPES motifs  
 FT /note= "cytoplasmic domain"  
 FT 808..813  
 FT /note= "used to design PCR oligos"  
 FT 870..875  
 FT /note= "used to design PCR oligos"  
 XX  
 PN W09519175-A1.  
 XX  
 XX 20-JUL-1995.  
 XX  
 XX 06-JAN-1995; 95WO-US00176.  
 XX  
 XX 14-JAN-1994; 94US-0183211.  
 XX  
 XX (UYJO ) UNIV JOHNS HOPKINS.  
 PA (UYPE-) UNIV PENNSYLVANIA.  
 XX  
 XX Civin CI, Gewirtz AM, Small D;  
 XX  
 XX WPI, 1995-263709/34.  
 DR N-PSDB; AAQ91356.  
 XX  
 XX Artificial STK-1 gene and gene-specific anti-sense oligo:nucleotide  
 PT used to treat neoplastic diseases and as bone marrow purging agents for  
 PT treating leukaemia and neoplasma  
 XX  
 PS Disclosure; Fig 1; 66pp; English.  
 XX  
 CC The STK-1 gene encodes a receptor PTK which is expressed in  
 CC proliferating hematopoietic stem cells but not in quiescent stem  
 CC cells. The STK-1 gene is also expressed in certain malignant cells  
 CC of non-hematopoietic origin. An antisense oligo specific for STK-1  
 CC is an oligo having a sequence (i) capable of forming a stable  
 CC triplex duplex with a portion of the STK-1 gene, or (ii) capable of forming  
 CC a stable duplex with a portion of an mRNA transcript of the STK-1  
 CC gene. Antisense oligos capable of forming a stable duplex with a  
 CC portion of a STK-1 mRNA transcript are given in AAQ91536 PT and in  
 CC AAQ91537 and AAQ91538. The antisense oligos of the invention are useful  
 CC in the treatment of hematologic malignancies characterised by  
 CC STK-1 expression. Several of the conserved domains of PTKs  
 CC including the ATP binding domain and the WMAPES motifs are  
 CC found in the STK-1 protein (see AAR75961 PT).  
 CC (updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 993 AA;

Query Match 84.1%; Score 4436.5; DB 16; Length 993;  
 Best Local Similarity 84.2%; Pred. No. 0;  
 Matches 837; Conservative 57; Mismatches 89; Indels 11; Gaps 4;

QY 1 MRALQSRDRRLLLVLLVLSYMLETVTNQDLPVTKCVLI SHENNGSAGKPSVYMRGS 60  
 DB 1 MPALA-RDAGTVLLVVSFAIFGTITNQDLPVTKCVLI INKNDSSVGKSSYPVMS 59  
 QY 61 PEDLQTPRRSGTVYEAATVEVASGSITLQVLAATPGDLSCLVFKHSLGCGPHD 120  
 DB 60 PEDLGCALRPQSSGTVYEAATVEVASGSITLQVLDVAPGNISCLVFKHSLGCGPHD 119  
 QY 121 LQNRGVSMALNVTTQAGEYLLHQTOSERANTVLTAVNRDQLVLRPVRKMEQ 180  
 DB 120 LQNRGVSMVILKMTQTAGEXYLLFQSEATNTYLTFTVSRITLTYLRPVRKMEQ 179

QY 181 DALLCISEGVPEPTVEWVLCSSHRESCKEGPAVVKKEKVLHELFGTDIRCCARNALGR 240  
 DB 180 DALVCISESVPEPTVEWVLCSSHRESCKEGPAVVKKEKVLHELFGTDIRCCARNALGR 239  
 QY 241 ECTKLTIDLNOAPOSTLPOLFLKGEPLWIRKAIHVNHGFLGTWELEKALEESYFE 300  
 DB 240 ECTRLFTIDLNOTPTTLPOLFLKGEPLWIRKAVVNHGFLGTWELEKALEEGNYFE 299  
 QY 301 MSTYSTNRTMIRILLAFVSSVGRNDGYTTCSSSKHPSQSALVTILEKGFINTATSSOEY 360  
 DB 300 MSTYSTNRTMIRILPAFVSSVARNDGYTTCSSSKHPSQSALVTIVEKGFINTATSSSEY 359  
 QY 361 EIDPYEKFCFSVRFKAYPRIRCTWIFSOASFCEQGLEGDGYISIKFCDHKNRGEVIFY 420  
 DB 360 EIDQYEEFCFSVRFKAYPQIRCTWTFSRKSPFCEQKGLDNGYSISKFCNHKQPGYIFH 419  
 QY 421 AENDDAQFTKMTNIRKPOVLANASASQASCSDDGYPLPSWTWKCKSDKSPNCTEIP 480  
 DB 420 AENDDAQFTKMTNIRKPOVLAEASASQASCSDDGYPLPSWTWKCKSDKSPNCTEIT 479  
 QY 481 EGVNKKANRKFVGQWVSSSTLNMSEAGKLLVKCCAYNSMGTSCTIFLNSPGPPFIQ 540  
 DB 480 EGVNKRANRKFVGQWVSSSTLNMSEAIKGFVKCCAYNSLGTSCETILLNSPGPPFIQ 539  
 QY 541 DNISFYATIGLCPLFIWLVILICHKKYKQFRYESOLOMLOVTPLDNEYFYVDFRDY 600  
 DB 540 DNISFYATIGVCLLFIWLVTLICHKKYKQFRYESOLOMVOVTVGSSDNEYFYVDFREY 599  
 QY 601 DLKWEPPRENLEFGKVLGSGAFGRVNNATAYGTSKTVSIQAVVXMKLKEKADSCKEKALM 660  
 DB 600 DLKWEPPRENLEFGKVLGSGAFGKVVNNATAYGTSKTVSIQAVVXMKLKEKADSSERELM 659  
 QY 661 SELKMMTHLGHDNI VNLIGACTLSGPVYLIPYCCYGDLLNLYLSKREKPHRTWTIFK 720  
 DB 660 SELKMMTQLGSHENI VNLIGACTLSGPVYLIPYCCYGDLLNLYLSKREKPHRTWTIFK 719  
 QY 721 EHNFSYPTFOAHNSNMFSGREVQLHPDLQSLGFGNSIHSDEDEIYEYENQKRLAEE 780  
 DB 720 EHNFSYPTFOGHPNSMFGREVQIHPDSQISLHGNFSHSEDEIYEYENQKRLAEE 777  
 QY 781 EDLNVLTFEEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGVVVKICDFGLARDIL 840  
 DB 778 EDLNVLTFEEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGVVVKICDFGLARDIM 837  
 QY 841 SDSSYVVRGNARLPVKWMAPESLFEGITYIKSDVMSYGILLWEIFSLGVNPPYGPV 900  
 DB 838 SDSSYVVRGNARLPVKWMAPESLFEGITYIKSDVMSYGILLWEIFSLGVNPPYGPV 897  
 QY 901 FYKLIOGPRMEOPFVATGIYFVMOSCAFDGRKRPSPNLTSLFLGCOLAEEAC--- 957  
 DB 898 FYKLIOGPRMDQPFYATEEIIYIMOSCAFDGRKRPSPNLTSLFLGCOLAEEAMQY 957  
 QY 958 ----IRTSIHLPKQAAPQORG-CLRAQSPORQVK 986  
 DB 958 VDRVSECPHTYQNRPPFGREMDLGLLSPQAQVE 991

RESULT 14  
 AAR67816  
 ID AAR67816 standard; Protein; 993 AA.  
 XX  
 AAR67816;  
 AC AC  
 XX XX  
 DT 25-MAR-2003 (updated)  
 DX 18-AUG-1995 (first entry)  
 XX  
 DE Flk2 receptor protein-tyrosine-kinase.  
 XX Human Flk2; receptor protein-tyrosine-kinase; primitive  
 KW hematopoietic cell; fetal liver kinase; diagnostic ligand  
 KW isolation; bone marrow disease therapy.  
 XX Homo sapiens.  
 OS

XX	Key	Location/Qualifiers	
FT	Peptide	1..27	
FT	Protein	/note= "signal peptide"	
FT	Domain	28..993	
FT	Domain	/note= "mature protein"	
FT	Domain	28..543	
FT	Domain	/note= "extracellular receptor domain"	
FT	Domain	544..563	
FT	Domain	/note= "transmembrane region"	
FT	Domain	564..993	
FT	Domain	/note= "intracellular catalytic domain"	
XX	WO9500554-A2.		
XX	05-JAN-1995.		
XX	17-JUN-1994;	94WO-US06944.	
XX	18-JUN-1993;	93US-0080244.	
PR	21-JUN-1993;	93US-0081508.	
PR	23-NOV-1993;	93US-0157490.	
XX	(UYPR-) UNIV PRINCETON.		
PA	Lemischka IR;		
XX	WFI; 1995-052014/07.		
DR	N-PSDB; AAQ81013.		
XX	Ligand for receptor protein tyrosine kinase - useful for the stimulation of primitive haematopoietic stem cells causing proliferation and/or differentiation		
PS	Disclosure; Fig 1b; 131pp; English.		
XX	The sequence corresponds to a human Flk2 (fetal liver kinase) receptor protein-tyrosine-kinase, which is expressed in primitive hematopoietic cells but not in mature hematopoietic cells. The protein is useful in isolation of receptor ligands, which have applications in diagnosis of bone marrow disorders and in stimulating proliferation and/or differentiation of primitive hematopoietic stem cells.		
CC	(Updated on 25-MAR-2003 to correct PN field.)		
XX	Sequence 993 AA;		
XX	Query Match 84.1%; Score 4429.5; DB 16; Length 993;		
XX	Best Local Similarity 84.1%; Pred. No. 0;		
XX	Matches 836; Conservative 57; Mismatches 90; Indels 11; Gaps 4;		
QY	1 MRALQSRDRRLLLVLSVMIETVTNQDLPVVKVLIHSHENNGSSAGKPSYRMVRGS 60		
Db	1 MPALA-RDAGTVLLVVFSAIFGTITNQDLPVVKVLIHSHENNGSSAGKPSYRMVRGS 59		
QY	61 PEDLQTPRQSGTVYEAATVEAAGSITLQVLTATPDGLSLVFKKSLCCQPHFD 120		
Db	60 PEDLGCALRPQSGTVYEAATVEAAGSITLQVLTATPDGLSLVFKKSLCCQPHFD 119		
QY	121 LQNRGVSMALNVTTQAGEYLLIHQSERANYTLFTVAVRDTQVLTQVLRPRPKMENQ 180		
Db	120 LQNRGVSMVILKMTQAGEYLLFIQSEATNYTLFTVAVRDTQVLTQVLRPRPKMENQ 179		
QY	181 DALLCISEGVPEPTVEVVLCSHRESCKECPAVVRKEEVLHFGTDIRCCARNALGR 240		
Db	180 DALVCISEGVPEPTVEVVLCSHRESCKECPAVVRKEEVLHFGTDIRCCARNALGR 239		
QY	241 ECTKLTIDLNQAPSTLPOLFVKVGEPLWIRKCAIHVNHGFGLTWELEKALEEGSYFE 300		
Db	240 ECTRLTIDLNQAPSTLPOLFVKVGEPLWIRKCAIHVNHGFGLTWELEKALEEGSYFE 299		
QY	301 MSTYSTNRTMIRILLAFVSSVGRNDTCYTCSSSKHPSQSALVTIIVGKGFINATNSSDEY 360		

Db	300	MSTYSTNRTMIRILLAFVSSVGRNDTCYTCSSSKHPSQSALVTIIVGKGFINATNSSDEY 359
QY	361	EIDPYEKECFSVRFKAYPRIRCTWTFPSQASPCBQRLGDEGYSISKPCDHKKQPGYIFY 420
Db	360	EIDQYEEFCFSVRFKAYPQIRCTWTFPSQASPCBQRLGDEGYSISKPCDHKKQPGYIFY 419
QY	421	AENDDAQFTKMTNIRKKPOVLANASASQSSDGYPLPSWTWKKCKSDKSPNCTEETIP 480
Db	420	AENDDAQFTKMTNIRKKPOVLANASASQSSDGYPLPSWTWKKCKSDKSPNCTEETIP 479
QY	481	EGVWKKANRKFVFGOWSSSTLNMSSEAGKGLLVKCCAYNSMGTSCTETIFLNSPGPPFIQ 540
Db	480	EGVWKKANRKFVFGOWSSSTLNMSSEAGKGLLVKCCAYNSMGTSCTETIFLNSPGPPFIQ 539
QY	541	DNISFYATIGLCLPPIVVLIVLICHYKKQFRYESOLOMIQVTPGLDNEFYFYVDPREY 600
Db	540	DNISFYATIGLCLPPIVVLIVLICHYKKQFRYESOLOMIQVTPGLDNEFYFYVDPREY 599
QY	601	DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISTGVSIQVAVKMLKEKADSCKEALM 660
Db	600	DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISTGVSIQVAVKMLKEKADSCKEALM 659
QY	661	SELKQMTLGHHDNIVNLLGACTLGGPIYLFYCCYGDLLNLYRSKREKPHRTWTEIFK 720
Db	660	SELKQMTLGHHDNIVNLLGACTLGGPIYLFYCCYGDLLNLYRSKREKPHRTWTEIFK 719
QY	721	EHNFSSYPTFOAHNSNMPGSRREVOLHPPLDQSGFNCSHSEDEIEYENOKRLAESEE 780
Db	720	EHNFSSYPTFOAHNSNMPGSRREVOLHPPLDQSGFNCSHSEDEIEYENOKRLAESEE 777
QY	781	EDNLVLPEDLLCFAYQVAKGMEFLPKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840
Db	778	EDNLVLPEDLLCFAYQVAKGMEFLPKSCVHRDLAARNVLVTHGKVVKICDFGLARDIM 837
QY	841	SDSYVVRGNARLPVKWMAPELFEIYTIKSDVMSYGLLWEIIFSLGVNPPGIPVDAN 900
Db	838	SDSYVVRGNARLPVKWMAPELFEIYTIKSDVMSYGLLWEIIFSLGVNPPGIPVDAN 897
QY	901	FYKLIQSGFKMOPFYATGFIYVMSQWAFDSRKRPSFNPILTSFLGQLAEEBAC--- 957
Db	898	FYKLIQSGFKMOPFYATGFIYVMSQWAFDSRKRPSFNPILTSFLGQLAEEBAC--- 957
QY	958	----IRTSIHLPKQAAPQORG-GLRAQSPQKQV 986
Db	958	VDGRVSECPHTYQNRPRPFSREMDLGLLSPQAQVE 991
XX	RESULT 15	
XX	AAR67536	
XX	ID	AAR67536 standard; Protein; 993 AA.
XX	AC	AAR67536;
XX	XX	
DT	25-MAR-2003	(updated)
DT	04-JUL-1995	(first entry)
XX	Human flk-2.	
XX	Fetal liver kinase-2; flk-2; protein tyrosine-kinase receptor;	
XX	hematopoiesis; stem cell.	
XX	Homo sapiens.	
XX	Key	Location/Qualifiers
FT	Peptide	1..27
FT	Domain	/label= Sig_peptide
FT	Region	28..543
FT	Region	/label= Extracellular_receptor_domain
FT	Region	544..563
FT	Region	/label= Transmembrane_region
FT	Domain	564..993
FT	Domain	/label= Intracellular_catalytic_domain
XX		

PN US5367057-A.  
FD 22-NOV-1994.  
XX  
XX  
XX 30-APR-1993; 93US-0055269.  
XX  
XX 02-APR-1991; 91US-0679666.  
XX 28-JUN-1991; 91US-0728913.  
XX 15-NOV-1991; 91US-0793065.  
XX 24-DEC-1991; 91US-0813593.  
XX 26-JUN-1992; 92US-0906397.  
XX 12-NOV-1992; 92US-0975049.  
XX 19-NOV-1992; 92US-0977451.  
XX 30-APR-1993; 93US-0055269.  
XX  
XX (UYPR-) UNIV PRINCETON.  
XX  
XX Lemischka IR;  
XX  
XX WPI, 1995-005894/01.  
XX N-PSDB; AAQ79069.  
XX  
XX Murine flk-2 receptor protein tyrosine kinase - used to stimulate  
XX proliferation and/or stimulation of primitive mammalian  
XX haematopoietic stem cells in vitro or in vivo.  
XX  
XX Disclosure; Fig. 2A-1F; 69pp; English.  
XX  
XX cDNAs encoding receptor protein tyrosine-kinases, mouse fetal liver  
XX kinase-2 (flk-2), human flk-2 and mouse flk-1 are given in AAQ79068-70,  
XX respectively, and the deduced amino acid sequences in AAQ67535-37,  
XX respectively.  
XX (Updated on 25-MAR-2003 to correct PF field.)  
XX  
XX Sequence 993 AA;  
XX  
Query Match 84.1%; Score 4429.5; DB 16; Length 993;  
Best Local Similarity 84.1%; Pred. No. 0;  
Matches 836; Conservative 57; Mismatches 90; Indels 11; Gaps 4;  
QY 1 MRALAQSDRLLLVLSVMILEVTNODLPVKVLIHSHENNGSSAGKPSYRMVRGS 60  
DB 1 MPALA-RDAGTVPLLVFSAIFGTITNODLPVKVLIHSHENNGSSAGKPSYRMVRGS 59  
QY 61 PEDLOCTPRQSEGTVEAATVEAEGSITLQVLAATPGDLSCLVFVKHSSLGCPHF 120  
DB 60 PEDLGCALRPOSCTVTEAAAVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCQPHF 119  
QY 121 LQNRGIVSMALNVTETQAGEYILHIQSERANYTVLFTVNVDTQLYVLRPFRKMNQ 180  
DB 120 LQNRGVSMVILKMTETQAGEYLLFIQSEATNYTILFTVNTLLYTLRRPFRKMNQ 179  
QY 181 DALLCISEGVPEPTVEWVLCSSHRESCKECPAVRKEEKVHELFGTDIRCCARNALGR 240  
DB 180 DALVCISESVEPEPVEWVLCDSQESCKEESPAVVKKEKVLHELFGTDIRCCARNELGR 239  
QY 241 ECTKLTIDLNAQPSLTLPQLFLKVGELWIRCKAIHVNHGFLTWELEDKALBEGSYFE 300  
DB 240 ECTELFTIDLNQPTTLTLPQLFLKVGELWIRCKAVHVNHGFLTWELENKALBEGSYFE 299  
QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGTYTSSSKHPSQSALVTILEKGFNATSSOEY 360  
DB 300 MSTYSTNRTMIRILLAFVSSVARNDTGTYTSSSKHPSQSALVTIVKGFNATSSSEY 359  
QY 361 EIDPYEFCFSVRKAYPRICRTWIFSOAPCORGLEDGYSIKPCDHKNKGEYIF 420  
DB 360 EIDQYEFECFSVRKAYPQIRCTWTFKSPCEQGLDNGYSISKPCNHKHQGEYIF 419  
QY 421 AENDDAQFTKMTNIRKKPQVLANASQASCSDDGYPLPSTWTKKCSKSPNCTBEIP 480  
DB 420 AENDDAQFTKMTNIRKKPQVLAEASQASCSFSDGYPLPSTWTKKCSKSPNCTBEIT 479  
QY 481 EGWVNRKANRKFQGWSSSTLNNSEAKGLLVKCCAYNSMGTSCETIFLNSPGFPFIQ 540

Db 480 EGWVNRKANRKFQGWSSSTLNNSEAKGLLVKCCAYNSMGTSCETIFLNSPGFPFIQ 539  
QY 541 DNISFYATIGLCLEPFIIVLIVLICHKYKQPRYESQLOMIQVTGFLDNEYFYVDFRDY 600  
Db 540 DNISFYATIGVCLLFIIVLITLICHKYKQPRYESQLOMQVVTGSSDNEYFYVDFREY 599  
QY 601 DLKWEFPRENLEFQKVLGSGAFGRVMNATAYIGISKTGVSIOAVVMKLEKADSCKEALM 660  
Db 600 DLKWEFPRENLEFQKVLGSGAFGRVMNATAYIGISKTGVSIOAVVMKLEKADSSREALM 659  
QY 661 SELKMTLGHHDNI VNLGACTLSGPVYLIFCYCCYGDLLNLYRSKREKHTWTETPK 720  
Db 660 SELKMTQLGSHENI VNLGACTLSGPVYLIFCYCCYGDLLNLYRSKREKHTWTETPK 719  
QY 721 EHNFSYPTFOAHNSNMPGSRREVQLHPPLDQLSGFNGNSIHSDEIEYENOKRLAEBEE 780  
Db 720 EHNFSYPTFOAHNSNMPGSRREVQLHPPLDQLSGFNGNSIHSDEIEYENOKRLAEBEE 777  
QY 781 EDNLVLTFBELLCPAYQVAKGMEFLEFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840  
Db 778 EDNLVLTFBELLCPAYQVAKGMEFLEFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIM 837  
QY 841 SDSSVVRGNARLPVKWMAPELSEFEGYTIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900  
Db 838 SDSSVVRGNARLPVKWMAPELSEFEGYTIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 897  
QY 901 FYKLIQSGFKMEQPFYATEGIYFVMSQWAFDSRKRPSFPNLTSLGCOLAEAEAC--- 957  
Db 898 FYKLIQNGFKMDQPFYATEEYIIMQSCWAFDSRKRPSFPNLTSLGCOLADAEEMMYQN 957  
QY 958 ----IRTSIHLPKQAAPQORG-GLRAQSPQSQVK 986  
Db 958 VQGRVSECPHTYQNRPFPSREMDLGLLSPOAQVE 991

Search completed: August 26, 2003, 07:24:09  
Job time : 67.4685 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2003, 07:27:59 ; Search time 45,4771 Seconds  
(without alignments)  
2878.506 Million cell updates/sec

Title: US-09-919-408A-2  
Perfect score: 5264  
Sequence: 1 MRALQRSERLLLVLSV.....RGGLRAQSPQVKIHRS 992

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 497079 seqs, 131961718 residues

Total number of hits satisfying chosen parameters: 497079

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :	Published Applications AA:*
1:	/cgn2_6/ptodata/1/pubaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/1/pubaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/1/pubaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/1/pubaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/1/pubaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/1/pubaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/1/pubaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/1/pubaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/1/pubaa/US09A_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/1/pubaa/US09B_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/1/pubaa/US09C_PUBCOMB.pep.*
12:	/cgn2_6/ptodata/1/pubaa/US09_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/1/pubaa/US10A_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/1/pubaa/US10B_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/1/pubaa/US10C_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pep.*
17:	/cgn2_6/ptodata/1/pubaa/US60_NEW_PUB.pep.*
18:	/cgn2_6/ptodata/1/pubaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5264	100.0	992	9	US-09-919-408-2
2	5264	100.0	992	10	US-09-872-136-2
3	4429.5	84.1	993	9	US-09-919-408-4
4	4429.5	84.1	993	10	US-09-872-136-4
5	1223	23.2	975	15	US-10-192-867-2
6	1221.5	23.2	972	10	US-09-944-807-10
7	1196	22.7	976	14	US-10-099-895-32
8	1196	22.7	976	15	US-10-192-867-4
9	1152.5	21.9	1088	11	US-09-961-403-4
10	1152.5	21.9	1089	9	US-09-769-987-2
11	1152.5	21.9	1089	10	US-09-919-497-90
12	1152.5	21.9	1089	10	US-09-866-510-2
13	1150.5	21.9	1089	10	US-09-955-363-36
14	1149.5	21.8	1089	10	US-09-866-510-10
15	1148.5	21.8	1089	10	US-09-866-510-4

16	1147.5	21.8	1089	10	US-09-866-510-8
17	1146.5	21.8	1089	10	US-09-866-510-6
18	1078	20.5	1090	10	US-09-866-510-14
19	1078	20.5	1106	10	US-09-866-510-22
20	1078	20.5	1106	10	US-09-955-363-2
21	1074	20.4	1106	10	US-09-866-510-16
22	1073	20.4	1106	10	US-09-866-510-20
23	1072	20.4	1106	10	US-09-866-510-18
24	998.5	19.0	1338	12	US-10-021-660-84
25	998.5	19.0	1338	15	US-10-059-585-44
26	961.5	18.3	1356	15	US-10-022-939-2
27	961.5	18.3	1356	15	US-10-100-405A-2
28	960.5	18.2	1356	11	US-09-969-037-7
29	954.5	18.1	1298	10	US-09-982-610-33
30	954.5	18.1	1298	15	US-10-262-538-32
31	954.5	18.1	1354	15	US-10-262-538-30
32	954.5	18.1	1363	11	US-09-375-248-2
33	952.5	18.1	1368	15	US-10-105-901-34
34	952.5	18.1	1363	11	US-09-375-248-19
35	943.5	17.9	1362	15	US-10-105-901-33
36	941	17.9	386	9	US-09-939-754-6
37	941	17.9	386	10	US-09-939-832-6
38	941	17.9	386	10	US-09-939-833-6
39	937.5	17.8	1367	9	US-09-766-678-2
40	936.5	17.8	1363	15	US-10-105-901-32
41	926.5	17.6	1367	9	US-09-919-408-6
42	926.5	17.6	1367	10	US-09-872-136-6
43	862	16.4	367	9	US-09-939-754-9
44	862	16.4	367	10	US-09-939-832-9
45	862	16.4	367	10	US-09-939-833-9

ALIGNMENTS

RESULT 1

US-09-919-408-2  
; Sequence 2, Application US/09919408  
; Patent No. US20020072077A1  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ImClone Systems Incorporated  
; STREET: 180 Varick Street  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/919,408  
; FILING DATE: 31-Jul-2001  
; CLASSIFICATION: <Unknown>  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 07/977,451  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: US 07/906,397  
; FILING DATE: 26-JUN-1992  
; APPLICATION NUMBER: US PCT/US92/05401  
; FILING DATE: 26-JUN-1992  
; APPLICATION NUMBER: TW 81102961  
; FILING DATE: 15-APR-1992  
; APPLICATION NUMBER: US PCT/US92/02750  
; FILING DATE: 02-APR-1992  
; APPLICATION NUMBER: US 07/813,593  
; FILING DATE: 24-DEC-1991

APPLICATION NUMBER: US 07/793,065  
 FILING DATE: 15-NOV-1991  
 APPLICATION NUMBER: US 07/728,913  
 FILING DATE: 28-JUN-1991  
 APPLICATION NUMBER: US 07/679,666  
 FILING DATE: 02-APR-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Feit, Irving N.  
 REGISTRATION NUMBER: 28,601  
 REFERENCE/DOCKET NUMBER: LEM-3-7P  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 212-645-1405  
 TELEFAX: 212-645-2054  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 992 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
 US-09-919-408-2

Query Match 100.0%; Score 5264; DB 9; Length 992;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MRALAQSDRLLLVLSVMILETVNODLPVTKCVLISHENNGSSAGKPSYRMVGRS	60
Db	1	MRALAQSDRLLLVLSVMILETVNODLPVTKCVLISHENNGSSAGKPSYRMVGRS	60
Qy	61	PEDLQCTPRQSEGVYEATVEAESGSIITVQLATPGDLSCLWFKHSSLCGCPHFD	120
Db	61	PEDLQCTPRQSEGVYEATVEAESGSIITVQLATPGDLSCLWFKHSSLCGCPHFD	120
Qy	121	LQNRGIVSMALNVTQAGEYLHIQSERANYTVLFTVNVDRDQLVLRPPYRKMENQ	180
Db	121	LQNRGIVSMALNVTQAGEYLHIQSERANYTVLFTVNVDRDQLVLRPPYRKMENQ	180
Qy	181	DALLCISEGPEPTVEVWLCSSHRESCKECPAVVRKEEVLHLEFGTDIRCCARNALGR	240
Db	181	DALLCISEGPEPTVEVWLCSSHRESCKECPAVVRKEEVLHLEFGTDIRCCARNALGR	240
Qy	241	ECTKLFIDLNQAPQSTLPQLFKVGEPLMIRCKAIHVNHGFGLTWELEDKALEEGSYFE	300
Db	241	ECTKLFIDLNQAPQSTLPQLFKVGEPLMIRCKAIHVNHGFGLTWELEDKALEEGSYFE	300
Qy	301	MSTYSTNRTMIRILLAFVSSVGRNDTGYTTCSSSKHPSQSALVTILEKGFNATSSQEEY	360
Db	301	MSTYSTNRTMIRILLAFVSSVGRNDTGYTTCSSSKHPSQSALVTILEKGFNATSSQEEY	360
Qy	361	EIDPYEKFCFSVRKAYPRIRCTWIFSQASPPCEQRGLEDGYSISKCDHKHKGPEYIFY	420
Db	361	EIDPYEKFCFSVRKAYPRIRCTWIFSQASPPCEQRGLEDGYSISKCDHKHKGPEYIFY	420
Qy	421	AENDDAQFTKMFTLNIRKKQVLNANASASQSDGYPLPSWTWKKCSKSPNCTEIP	480
Db	421	AENDDAQFTKMFTLNIRKKQVLNANASASQSDGYPLPSWTWKKCSKSPNCTEIP	480
Qy	481	EGVWKKANKRVFGQWSSSTLNMSKAGLLVKCCAYNSMGTSCTETFLNSPGPPFFIQ	540
Db	481	EGVWKKANKRVFGQWSSSTLNMSKAGLLVKCCAYNSMGTSCTETFLNSPGPPFFIQ	540
Qy	541	DNISFYATIGLCIPFIVLVILCHYKQFRYESQLQMTQVTPDLNNEYFYVDFRDEY	600
Db	541	DNISFYATIGLCIPFIVLVILCHYKQFRYESQLQMTQVTPDLNNEYFYVDFRDEY	600
Qy	601	DLKWEPPRENLEFGKVLGSAFGVMNATYIGISKTGVSQVAVKMLKEKADSCKEALM	660
Db	601	DLKWEPPRENLEFGKVLGSAFGVMNATYIGISKTGVSQVAVKMLKEKADSCKEALM	660
Qy	661	SELKMWTHLGHNDIVNLLGACTLSGPFVLIFFCYCGDLLNLYRSKREKPHRTWTIFK	720
Db	661	SELKMWTHLGHNDIVNLLGACTLSGPFVLIFFCYCGDLLNLYRSKREKPHRTWTIFK	720

RESULT 2  
 US-09-872-136-2  
 ; Sequence 2, Application US/09872136  
 ; Patent No. US20020119545A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lemischka, Ihor R.  
 ; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
 ; RECEPTORS AND THEIR LIGANDS  
 ; NUMBER OF SEQUENCES: 10  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Inclone Systems Incorporated  
 ; STREET: 180 Varick Street  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10014  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/872,136  
 ; FILING DATE: 01-Jun-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/208,786  
 ; FILING DATE: <Unknown>  
 ; APPLICATION NUMBER: US/09/021,324  
 ; FILING DATE: <Unknown>  
 ; APPLICATION NUMBER: US/07/977,451  
 ; FILING DATE: 1992-11-19  
 ; APPLICATION NUMBER: US 07/906,397  
 ; FILING DATE: 26-JUN-1992  
 ; APPLICATION NUMBER: US PCT/US92/05401  
 ; FILING DATE: 26-JUN-1992  
 ; APPLICATION NUMBER: TW 81102961  
 ; FILING DATE: 15-APR-1992  
 ; APPLICATION NUMBER: US PCT/US92/02750  
 ; FILING DATE: 02-APR-1992  
 ; APPLICATION NUMBER: US 07/813,593  
 ; FILING DATE: 24-DEC-1991  
 ; APPLICATION NUMBER: US 07/793,065  
 ; FILING DATE: 15-NOV-1991  
 ; APPLICATION NUMBER: US 07/728,913  
 ; APPLICATION NUMBER: US 07/679,666  
 ; FILING DATE: 02-APR-1991  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Feit, Irving N.  
 ; REGISTRATION NUMBER: 28,601

Qy	721	EHNFSYPTFOAHNSNMPGSRVQLHPPDLQSLGFGNSIHSDEIEYENOKRLAESEE	780
Db	721	EHNFSYPTFOAHNSNMPGSRVQLHPPDLQSLGFGNSIHSDEIEYENOKRLAESEE	780
Qy	781	EDNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL	840
Db	781	EDNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL	840
Qy	841	SDSSYVVRGNARLPVKWMAPESLPEGIYTKSDVWSYGILLWEIFSLGWNYPGIPVDAN	900
Db	841	SDSSYVVRGNARLPVKWMAPESLPEGIYTKSDVWSYGILLWEIFSLGWNYPGIPVDAN	900
Qy	901	FYKLIQSGFMQEPFYATEGIFYFMQSCWAFDSKRPSFPNLTGFLGQLAEEBACIRT	960
Db	901	FYKLIQSGFMQEPFYATEGIFYFMQSCWAFDSKRPSFPNLTGFLGQLAEEBACIRT	960
Qy	961	SIHLPKQAAPOQRGLRAQSPQRQVKIHRRS	992
Db	961	SIHLPKQAAPOQRGLRAQSPQRQVKIHRRS	992

REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 992 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-872-136-2

Query Match 100.0%; Score 5264; DB 10; Length 992;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALQSDRRLLLVLSVMILETVTNQDLVPIKCVLISHENNGSSACKPSSYMRVGS 60  
DB 1 MRALQSDRRLLLVLSVMILETVTNQDLVPIKCVLISHENNGSSACKPSSYMRVGS 60  
QY 61 PEDLOCTPRROSEGVYEAATVEAESGSIITIQVQLATPGDLSCLWPKHSLGQCPHFD 120  
DB 61 PEDLOCTPRROSEGVYEAATVEAESGSIITIQVQLATPGDLSCLWPKHSLGQCPHFD 120  
QY 121 LQNRGIVSMAILNVETQAGEYLLHIQSERANTVLTFTVNVVDTQLYVLRPFYFRMENQ 180  
DB 121 LQNRGIVSMAILNVETQAGEYLLHIQSERANTVLTFTVNVVDTQLYVLRPFYFRMENQ 180  
QY 181 DALLCISEGVPEPTVEWVLCSSHRSCKEGBPAVVRKEKVLHELPGTDIRCCARNALGR 240  
DB 181 DALLCISEGVPEPTVEWVLCSSHRSCKEGBPAVVRKEKVLHELPGTDIRCCARNALGR 240  
QY 241 ECTKLFITDLNQAPOSTLPOLFLKVGEPWIRCKAIHVNHGGLTWELDKALEGSYFE 300  
DB 241 ECTKLFITDLNQAPOSTLPOLFLKVGEPWIRCKAIHVNHGGLTWELDKALEGSYFE 300  
QY 301 MSTYSTNTRMIRILLAFVSSVGRNDGYTSCSSKHPQSALVTILEKGFINATSQEY 360  
DB 301 MSTYSTNTRMIRILLAFVSSVGRNDGYTSCSSKHPQSALVTILEKGFINATSQEY 360  
QY 361 EIDPYEKFCFSVRKAYPRIRCTWIFSQAFFCEQRLGLEDGYSISKFDHKNKPGYIFY 420  
DB 361 EIDPYEKFCFSVRKAYPRIRCTWIFSQAFFCEQRLGLEDGYSISKFDHKNKPGYIFY 420  
QY 421 AENDDAQFTKMTNIRKPKQVLANASASQASCSDDGYPLPSWTWKKCSDKSPNCTEIP 480  
DB 421 AENDDAQFTKMTNIRKPKQVLANASASQASCSDDGYPLPSWTWKKCSDKSPNCTEIP 480  
QY 481 EGVNKKANRKYFGQVSSSTLNMSAGKGLLVKCCAYNSMGTSCTRTIPLNSGPPPPFIQ 540  
DB 481 EGVNKKANRKYFGQVSSSTLNMSAGKGLLVKCCAYNSMGTSCTRTIPLNSGPPPPFIQ 540  
QY 541 DNISFYATIGLCLPFIIVLVLVILCHYKQFRYESQLQMIQVTPLDNBYFYVDPDY 600  
DB 541 DNISFYATIGLCLPFIIVLVLVILCHYKQFRYESQLQMIQVTPLDNBYFYVDPDY 600  
QY 601 DLKWFPPRENLEFGKVLGSGAFGRVNNATAYGISTGVSIOQAVMLKEKADSCKEALM 660  
DB 601 DLKWFPPRENLEFGKVLGSGAFGRVNNATAYGISTGVSIOQAVMLKEKADSCKEALM 660  
QY 661 SELKMWTHLGHNDINVLGACTLSPGVYLIPEYCCVGDLLNVLRSKREKFRHTWTEIPK 720  
DB 661 SELKMWTHLGHNDINVLGACTLSPGVYLIPEYCCVGDLLNVLRSKREKFRHTWTEIPK 720  
QY 721 ENHFSYPTFOAHNSMPCGSRVQLHPDQLSGFNGNSIHSEDEIEYENQRLAESEE 780  
DB 721 ENHFSYPTFOAHNSMPCGSRVQLHPDQLSGFNGNSIHSEDEIEYENQRLAESEE 780  
QY 781 EDNLNLTTPEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGKVKITCDFGLARDIL 840  
DB 781 EDNLNLTTPEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGKVKITCDFGLARDIL 840

QY 841 SDSSYVRGNARLPVKWNPESLPEGIYTIKSDVMSYGILLWEIFSLGWNPIPGIPVDAN 900  
DB 841 SDSSYVRGNARLPVKWNPESLPEGIYTIKSDVMSYGILLWEIFSLGWNPIPGIPVDAN 900  
QY 901 FYKLQSGFKMEQPFYATEGIYFVWQSWAFDSRRKRPSPFNLTSLGCOLAEAEACIRT 960  
DB 901 FYKLQSGFKMEQPFYATEGIYFVWQSWAFDSRRKRPSPFNLTSLGCOLAEAEACIRT 960  
QY 961 SIHLPKQAAPQORGGLRAQSPQRQVKIHRERS 992  
DB 961 SIHLPKQAAPQORGGLRAQSPQRQVKIHRERS 992

RESULT 3  
US-09-919-408-4  
Sequence 4, Application US/09919408  
Patent No. US2002007207A1  
GENERAL INFORMATION:  
APPLICANT: Lemischka, Ihor R.  
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
RECEPTORS AND THEIR LIGANDS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ImClone Systems Incorporated  
STREET: 180 Varick Street  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10014  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/919,408  
FILING DATE: 31-Jul-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/977,451  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US 07/906,397  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: US PCT/US92/05401  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: TW 81102961  
FILING DATE: 15-APR-1992  
APPLICATION NUMBER: US PCT/US92/02750  
FILING DATE: 02-APR-1992  
APPLICATION NUMBER: US 07/813,593  
FILING DATE: 24-DEC-1991  
APPLICATION NUMBER: US 07/793,065  
FILING DATE: 15-NOV-1991  
APPLICATION NUMBER: US 07/728,913  
FILING DATE: 28-JUN-1991  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Felt, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 993 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-919-408-4



Query Match 84.1%; Score 4429.5; DB 9; Length 993;  
Best Local Similarity 84.1%; Pred. No. 0;  
Matches 836; Conservative 57; Mismatches 90; Indels 11;

QY		1	MRALAQBSDRLLILLVLSVMILEFTVNODLPVIKCVLISHENNNGSSAGKSPSSRYMVRGS	60
Db		1	MPALA-RDAGTVPLVVVSAMIFGTITNQDDLFVIKCVLINHNNDSSVGKSSSTPMVSES	59
QY		61	PEDLQCTPRRQSEGVTEAATVEVAESGSI TLQVLATPGDLSC LWMFKHSLSGCQHFD	120
Db		60	PEDLGCA LR PQSSGTVYEAAAVEVDVSASITLQLVDPAGNISCLWFKHSSLNCQHFD	119
QY		121	LQNRGISMAILNVETOAGEYLHIHQSERANYTTLFTVNVRTDQLYVLRPYPRKMENQ	180
Db		120	LQNRGVSVMLKMTETQAGEYLLFIQSEANTYTILFVTISRNTLLXYLRPYPRKMENQ	179
QY		181	DALLCSIEGPPEPTVEUWLCCSHRESCKEEGPAVVRKEEKVLHELFGTDIRCCARNALGR	240
Db		180	DALVCISESVP EPIVENWLCD SQESCKEESP AVVKKEEKVLHELFGTDIRCCARNELGR	239
QY		241	ECTKLFTIDLNAQPQSTLPQLFLKVGEPLWRICKAIHVNHGGFLTWELEDKALBEGSYFE	300
Db		240	ECTRLFTIDLNQTPQTLLPOLFLKVGEP LWRICKAVHVNHGGFLTWELENKALBEGSYFE	299
QY		301	MSTYSTNRTMIRILLAFVSSVGRNDTGYYTCSSSKHPQSQALVILEKGFINATSSQBEEY	360
Db		300	MSTYSTNRTMIRILFAFVSSVARNDTGYYTCSSSKHPQSQALVTIVGKGFINATNSSBDY	359
QY		361	EIDPYEKFCFSVRPKAYPRICRTWI FSOASPBCORGLEDCYSISKFCDHKNKPGEYIFY	420
Db		360	EIQOYEFCEFSVRPKAYPQIRCTWTFRRKSPCEQKGLDN GYSISKFCNHKHQPGEYIFY	419
QY		421	AENDDAQTKMFTLNIRKKPOVLANASQASCSDGYPLPSWTWKCDKSPNCTEBIP	480
Db		420	AENDDAQTKMFTLNIRRKPOVLAEASQASCSDGYPLPSWTWKCDKSPNCTEBIT	479
QY		481	EGWNKXANKRVFQOWYSSSTLANNSEAGKLLLVKCCAYNSMGTSCTETFLNSPGPPPIQ	540
Db		480	EGWNNRXANKRVFQOWYSSSTLANNSEAIKGPLVRCCA YNSLGTSETILLNSPGPPFPPIQ	539
QY		541	DNISFYATIGLC LPFI VVULI CHKYKKQFRYESQLQMI QVTGPD NNEYFYVDFRDY EY	600
Db		540	DNISFYATI GVCLLFIVVULTLI CHKYKKQFRYESQLQ MWQVTGSSDNEYFYVDFREY EY	599
QY		601	DLKWEPRENLEFGKVLGSGAFGRVMNATAYGISKTGVSIOQAVKMLEKADSCKEKALM	660
Db		600	DLKWEPRENLEFGKVLGSGAFGKVMNATAYGISKTGVS IQAVAKMLEKADSSEREALM	659
QY		661	SELKMMTHLGHHDNVNLGACTLSGPVYLI FEYCCTGD LLYLRSKEKHFWTWTETFK	720
Db		660	SELKMMYQLGSHENI VNLLGACTLSGPTYLI FEYCCTGD LLYLRSKEKHFWTWTETFK	719
QY		721	EHNFSSYPTFOAHGNSSMPGSRREVQLPHLPDLQSGFN GNSIHSDEIEIYENOKRLABEEE	780
Db		720	EHNFSPYPTQSHFNSSMPGSRREVQIH PDSDOI SGLHGNSFHSEDEIEIYENQKRL-- BEE	777
QY		781	EDNLVLTPEDLLCFAYQVAKGMWFLEPKSCVHRDLAARNVLVTHGKVVKICDFGLGIARDIL	840
Db		778	EDNLVLTPEDLLCFAYQVAKGMWFLEPKSCVHRDLAARNVLVTHGKVVKICDFGLGIARDIM	837
QY		841	SDSNVYVRGNARLPVKMKAPESLPEGIYTIKSDVMSYGILLWEI FSLGVNYPGIPVDAN	900
Db		838	SDSNVYVRGNARLPVKMKAPESLPEGIYTIKSDVMSYGILLWEI FSLGVNYPGIPVDAN	897
QY		901	FYKLIQSGFKMEQPFYATEGITYFVWQSCWAFPSDRKR PSFPNLTSLPGCOLAEABEAC---	957
Db		898	FYKLIQNGFKMDQPFYATEEIIYIMQSCWAFPSDRKR PSFPNLTSLPGCOLADEAMEYQN	957
QY		958	----IRTSIHLPKQAAPQORG-GLRAQOSPQRQVK	986
Db		958	VDRGVSECPHTYQNRPPFSRMDLGLLSPQAQVE	991

```

RESULT 4
US-09-872-136--4
; Sequence 4, Application US/09872136
; Patent No. US20020119545A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOPIPOTENT HEMATOPOIETIC STEM CELL
; ; ; RECEPTORS AND THEIR LIGANDS
; ; ;
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Imclone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/872,136
; FILING DATE: 01-Jun-2001
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/208,786
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/09/021,324
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/07/977,451
; FILING DATE: 1992-11-19
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-7P
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
;
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 993 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-09-872-136--4
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```

61 PEDLQCTPRRQSEGTVEAATVEVAESGSITLQVQLATPGDLSCLWVFKHSSLGCQPHFD 120 QY

Db 60 PEDLGCALPQSSGTYEAAVEVDVSATITLQVLVDAPGNIISCLWFKHSSLNCPHF 119  
QY 121 LQNRGVSMALNVITETQAGEYLLHIOSEANVTYVLTNNVDTQLYVLRPFRKMEQ 180  
Db 120 LQNRGVSMVILKWTETQAGEYLLFQSEATNTILFTVSIRNTLLYLRPFRKMEQ 179  
QY 181 DALLCISEGPEPTVWVLCSSHRESCKEKGPAVRKEEVLHELFGTDIRCCARNALGR 240  
Db 180 DALVCISESPEPIVEWVLCDSQGESCKESPAPVVKKEKVLHELFGTDIRCCARNALGR 239  
QY 241 ECTKLTIDLNOAPOSTLPOLFLKVGEPWIRCKAHVNHGFLTWELEDKALEESYFE 300  
Db 240 ECTRLFTIDLNOPTQLPOLFLKVGEPWIRCKAHVNHGFLTWELEDKALEEGNYFE 299  
QY 301 MSTYSNTRMIRILLAFVSSVGRNDTGYTCSSSKHPSQSAVLTILEKGFINATSSOEY 360  
Db 300 MSTYSNTRMIRILFAFVSSVARNDTGYTCSSSKHPSQSAVLTIVCKGFINATSSDY 359  
QY 361 EIDPYKCFVSFRKAYPRIRCTWIFSQASFPCEQGLDGYISKPCDHKNKPGYIFY 420  
Db 360 EIDQYBEFCFSVRKAYPQIRCTWTSRKSFPCEQGLDNGYSISKPCNHKQPGYIFY 419  
QY 421 AENDDAQFTKMTFLNIRKPKQVLANASASQASCSGYPPLSPWTWKCKSKSPNCTEIP 480  
Db 420 AENDDAQFTKMTFLNIRKPKQVLANASASQASCSGYPPLSPWTWKCKSKSPNCTEIT 479  
QY 481 EGVNWKANRKYFGQVSSSTLNMBSAGKLLVKCCAYNSMGTSCTIFLNSPQPPFIQ 540  
Db 480 EGVNWKANRKYFGQVSSSTLNMBSAGKLLVKCCAYNSMGTSCTIFLNSPQPPFIQ 539  
QY 541 DNI SFYATIGLCLPFIWTLVILVILCHYKQFRYESQLQMIQVTPLDNEYFYVDFRDEY 600  
Db 540 DNI SFYATIGVCLLFIWTLVILVILCHYKQFRYESQLQMIQVTPLDNEYFYVDFREY 599  
QY 601 DLKWEPPRENLEFGKVGSGAFQVNMATAYGSKTGVSIQVAVMLKEKADSCKEALM 660  
Db 600 DLKWEPPRENLEFGKVGSGAFQVNMATAYGSKTGVSIQVAVMLKEKADSCSEREALM 659  
QY 661 SELKMTLGHGHNINVLGACTLSGPVILI FEYCCVGDLLNLYLSKREKFRHTWEIFK 720  
Db 660 SELKMTQLGSHENINVLGACTLSGPVILI FEYCCVGDLLNLYLSKREKFRHTWEIFK 719  
QY 721 EHNFSYPTFOAHNSMPSGSRVQLHPDLDQSGFNGNSIHSEDELEYENQKRLAESEE 780  
Db 720 EHNFSYPTFOAHNSMPSGSRVQLHPDLDQSGFNGNSIHSEDELEYENQKRLAESEE 777  
QY 781 EDNLVLTFFEDLLCFAYQVAKGMEFLPKSCVHRDLAARNVLTGHKVVYKICDFGLARDIL 840  
Db 778 EDNLVLTFFEDLLCFAYQVAKGMEFLPKSCVHRDLAARNVLTGHKVVYKICDFGLARDIM 837  
QY 841 SDSSYVVRGNARLPVKWMAPELFEIGYITIKSDVWSYGILLWEIFSLGVNYPGIPVDAN 900  
Db 838 SDSSYVVRGNARLPVKWMAPELFEIGYITIKSDVWSYGILLWEIFSLGVNYPGIPVDAN 897  
QY 901 FYKLIQGFMEOPFYFATEGIYFMQSCWAFDSKRPSPFNLTSLFLGCQLAEEAEAC--- 957  
Db 898 FYKLIQGFMEOPFYFATEGIYFMQSCWAFDSKRPSPFNLTSLFLGCQLAEEAEAMQY 957  
QY 958 ----IRTSIHLPKQAAAPQORG-GLRAQSPQRQVK 986  
Db 958 VDRGVSECPHTYQNRPRPFSRMDLGLLSPQAQVE 991

## RESULT 5

US-10-192-867-2  
; Sequence 2, Application US/10192867  
; Publication No. US20030084466A1  
; GENERAL INFORMATION:  
; APPLICANT: BLUME-JENSEN, Peter  
; APPLICANT: HUNTER, Tony  
; TITLE OF INVENTION: SYSTEM AND METHOD FOR CONTROLLING MALE FERTILITY  
; FILE REFERENCE: SALKINS.002C1

; CURRENT APPLICATION NUMBER: US/10/192,867  
; CURRENT FILING DATE: 2002-07-08  
; PRIOR APPLICATION NUMBER: US 60/175,625  
; 2000-01-11  
; PRIOR APPLICATION NUMBER: PCT/US01/00573  
; 2001-01-08  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 975  
; TYPE: PRT  
; ORGANISM: MUS MUSCULUS  
US-10-192-867-2

Query Match 23.2%; Score 1223; DB 15; Length 975;  
Best Local Similarity 32.1%; Pred. No. 5.9e-94;

Matches 324; Conservative 172; Mismatches 331; Indels 182; Gaps 39;

QY 36 CVLISHENNSSACKSSYMRVSGSPEDLOCTPRQSEGITVYEAATVEVABSGSITLQVQ 95  
Db 12 CVLLVLLRGQTATSPSA-----SPGEPPSIHPAQSELIVEAGD-TL--- 54  
QY 96 LATPGDLSCD-----WFKHSSLGCPHF-----DLQNRGIVSMALNVETQAGEYLL 144  
Db 55 -----SLTCLDPPFVWTFK-----TYFNEMVKNKNWIOKA-----PATRTGYTC 98  
QY 145 HIQSERANYTVLFTVNRD-TQYVLRPFRKMEQDALL-----CIS 187  
Db 99 ---SNSNGLTSSIVYFVRDPKFLVGLPLFGK-EDSDALVRCPLTDPQVSNYSILIECDG 154  
QY 188 EG-----VPEPTVWVLCSSHRESCKEKGPAVRKEEVLHELFGTDIRCCARNALGR 240  
Db 155 KSLPTDLTFVPNP-----KAGITIKNVKGAYH-----RLCVRCAQR 191  
QY 241 ECTKLT-----FTIDLNOA-----POSTLPOL--FLKVGEPWIRCKAHVNHGFLTW--- 286  
Db 192 DGTWLSHDKFTLVREAIKAI PVVSPETSHLLKGGDTFTVVCITKDVSTVSNMVLKQON 251  
QY 287 -----ELBDKALEGSEYFEMSTYSTNRMTIRILLAFVSSVGRNDTGYTCSSSKHPSQ 339  
Db 252 PQQHIAQVKNHWRGDF-----NYERQETLT-----ISSARVDDSGVFCYANNFTGS 301  
QY 340 SALVT-----ILEKGFINAT-SQBEYEIDPYEKFCSVRFKAYPR-IRCTWIFSQASFPCE 394  
Db 302 ANVTTLTKVKEGFINISPVKNVTFTVDGENVDLVVEYEAYPKPEHQOMIYMNRT--SA 359  
QY 395 QRGLE-----DGYISKPCDH-----KNKPGEYIFYAENDDAQFTKMTFLNIRKPKQVL 443  
Db 360 NKGDYVYKSDNKNIRYVNLRLTLKGTGGTYTFLVSNDSASASVTFNYYVNTKPEIL 419  
QY 444 A--NASASQASCSGSDGYPLSPWTWKCKSKSPNCTEII-PEGVNMKANRKYFGQVWSSS 500  
Db 420 TYDRLNGMLQCVABGPEPTIDWYFCTGABQRCCTTPVSPVDVQVQNVSPSPGKLVVQS 479  
QY 501 TLNMBAGKLLVKCCAYNSMGTSCTIFLNSPQPPFIQDNISFYATIGLCLPFIV--- 557  
Db 480 SIDSSVFRHNGTVECKASNDVGKS--SAFFN-----PAFKEIQOAHLTFTPLLIGFVVAAG 533  
QY 558 ---VLVLVILCHYKQFRYESQLQMI-QVTGPLDNEYFYVDFRDEYDLKWEPPRENLEF 613  
Db 534 AMGIIVNVLTYKYLOKPMYEVQWKVVEEING---NNYVYIDPTQLPDYDKWEFFRNRLSF 590  
QY 614 GKVLGSAFGVNMATAYGSKTGVSIQVAVMLKEKADSCKEKALMSELKMTLGHGHD 673  
Db 591 GKTLLGAGAFKVVEATAYGLIKSDAMTVAVKMLKPSAHLTERALMSELKVLVYLGNHM 650  
QY 674 NIVNLLGACTLSGVPVLI FEYCCYGDLLNLYLSKREKFR-----HRTWTEIFKHNFSY 727  
Db 651 NIVNLLGACTVGGTTLVITEYCCYGDLLNLFRRKDSFI FSKQEEQAEAAALYKNLLHSTE 710  
QY 728 PTOAHNSNM---PGSREVQLHPDLDQSGFNGNSIHSEDELEYENQKRLAESEEDLN 784  
Db 711 PSCDS-SNEYMDMKPGVSYV-VPTKTDK-----RRSARIDSYIERDVTPTAIMEDELALD 763



```

Db 109 VF---VRDPAKFLVDRSLYKEDNDTLVRCPLTDPVNTNYSKGCQKPLPKDLRFIPD 165
Qy 193 PTVEWLVCSHRE-----SKEEGPAVVRKEKVLHFLPGTDIRCCARNALGRECTK 244
Db 166 PKAGIMIKSVKRAVHRLCHLCSVDOEG-----KSVLSEKFLKVRPAFK-----A 210
Qy 245 LFTIDLNOAPOSTLPQLFKVGEPLWIRCKAHVNHGFLGTWELEDK--ALEE----- 295
Db 211 VPVSVSKA-----SYLLREGEETVCTIKDVSSVYSTWKRENSQTKLQEKYNMWH 264
Qy 296 GSYEMSTYTNRTMIRILLAFVSSVGRNDTGYTCSSKHPQSALVTILE---KGFIN 352
Db 265 GDF-----NYERQATLT-----ISSARVNDSGVFCYANNTFGSANVTTLLEWVDKGIN 314
Qy 353 A-TSSQBEYEIDPYEKFCSYRFAKAYPR-IRCTWIFSOASP--PCEORGLEGYGISKFC 408
Db 315 IFPMINTTVFVNDGENVDLIVEYEAFFPEHQOQWIMYRTFTDKWEDYPKSENEINRIV 374
Qy 409 DHKN-----KPEYIFYAENDDAQFTKMFNTLNIRKQVOLA--NASASQASCSDDGYP 459
Db 375 SELHLRLKGTGGTYTFLVNSDVNAALAFNVYNTKPEILTYDLRVNGLMOCVAAAGFP 434
Qy 460 LPSWTWKCSKSPNCTBEI--PEGVWKKANKRVFGQWVSSSTLNMSEAGKLLVKCCAY 518
Db 435 EPTIDWYFCPCTEQRCASVLPVDVQTLNSSGPPFGKLVQSSIDSSAFKHNGTVECKAY 494
Qy 519 NSMGTSCETIFLNSPGPPFPFIQDN-----ISFYATIGLCLPFIWLVILIC 564
Db 495 NDVGKT--SAYFN-----PAFKGNKEQIHPTLFTPLLIGFVIVAGM---MCIIVMILT 544
Qy 565 HKYKQKRYESQLOMI-QVTGPLDNEYFYVDRDYEDLKWEPFRENLEFGKVLGSAFG 623
Db 545 KYLQKPMYEVQWKVBEING---NNVYIIDPTQLPYDHKEFFPNRLSPFGKTLGAGAFG 601
Qy 624 RVNATAYGISTGVSIOQAVKMLKEKADSEKALMSELKMWTHLGHNDINVLGACT 683
Db 602 KVEATAYGLIKSDAAMTVAVKMLKPSAHLTEREALMSELKVLVSYLGNHMINVLGACT 661
Qy 684 LSGPVYLIFECYCGDILNLYLRSKREKPHRTWTETFEKHNFSSTYPTFOAHNSNMPGSR 743
Db 662 IGGPTLVITEYCCYGDLLNFLRRKRSFICSKQEDHAEALYKYNLLHSHKSSCSDSTNEY 721
Qy 744 VOLHPPDLQD---SGFNNGSIHSEDEIYENQKRLAEEEDLNVLTTFEDLLCPAYQVAK 800
Db 722 MDMPKGVSVVPTKADKRSVRIGSYIERDVTPTAIMEDEALDL---EDLLSFSYQVAK 778
Qy 801 GMEFLFKSCVHRDLAARNVLVTHGKVKIKCDFGLARDILSDSSVVRGNARLPVKWMAP 860
Db 779 GMAFLASKNCIHRDLAARNILLTHGRITKICDFGLARDIKNDSNVVVKGNARLPVKWMAP 838
Qy 861 ESLPEGIYTIKSDVWSGILLWEIFSLGVNYPYGPIDVANTFKLIQSOFKMEQPPYATEG 920
Db 839 ESIFNCYVTFESDVMYSIGIFLWELFSLGSSYPYGPMPVDSPKFKYMKIEGFRMLSPEHAPAE 898
Qy 921 IYFVMSQWAFDSRKPSFPNLTSLFGQLAE 953
Db 899 MYDLMTCWDADPLKRPFTFKQIVLIEKQISES 931

```

## RESULT 8

US-10-192-867-4

; Sequence 4, Application US/10192867

; Publication No. US20030084466A1

; GENERAL INFORMATION:

; APPLICANT: BLUME-JENSEN, Peter

; APPLICANT: HUNTER, Tony

; TITLE OF INVENTION: SYSTEM AND METHOD FOR CONTROLLING MALE FERTILITY

; FILE REFERENCE: SALKINS.002C1

; CURRENT APPLICATION NUMBER: US/10/192.867

; CURRENT FILING DATE: 2002-07-08

; PRIOR APPLICATION NUMBER: US 60/175,625

; 2000-01-11

; PRIOR APPLICATION NUMBER: PCT/US01/00573

; 2001-01-08

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 976

; TYPE: PRT

; ORGANISM: HOMO SAPIENS

US-10-192-867-4

Query Match 22.7%; Score 1196; DB 15; Length 976;

Best Local Similarity 31.4%; Pred. No. 1.1e-91;

Matches 312; Conservative 172; Mismatches 361; Indels 148; Gaps 32;

Qy 36 CVLISHENNGSSAGKPPSSYRMVRGSPEDLQCTPRQSEGTVEEATVEVAESGSILOVQ 95

Db 12 CVLLLLLRVQTSSQPS-----VSPGEPSPSIHPGKSDLIVRGD-EIRLL 57

Qy 96 LATPGDLSCLMVFKHSSLCQPHFDLQNRGIVSMALNVTETQAGEYLLHISERANYTV 155

Db 58 CTDPGFVK--WTFE---ILDETENKQNEWITEKA---EATNTGKYCTCNKHGLNSIY 108

Qy 156 LFTVNVDR-TOLYVLRPRPYFRKMNQDALLCI-----SEG-----VPE 192

Db 109 VF---VRDPAKFLVDRSLYKEDNDTLVRCPLTDPVNTNYSKGCQKPLPKDLRFIPD 165

Qy 193 PTVEWLVCSHRE-----SKEEGPAVVRKEKVLHFLPGTDIRCCARNALGRECTK 244

Db 166 PKAGIMIKSVKRAVHRLCHLCSVDOEG-----KSVLSEKFLKVRPAFK-----A 210

Qy 245 LFTIDLNOAPOSTLPQLFKVGEPLWIRCKAHVNHGFLGTWELEDK--ALEE----- 295

Db 211 VPVSVSKA-----SYLLREGEETVCTIKDVSSVYSTWKRENSQTKLQEKYNMWH 264

Qy 296 GSYEMSTYTNRTMIRILLAFVSSVGRNDTGYTCSSKHPQSALVTILE---KGFIN 352

Db 265 GDF-----NYERQATLT-----ISSARVNDSGVFCYANNTFGSANVTTLLEWVDKGIN 314

Qy 353 A-TSSQBEYEIDPYEKFCSYRFAKAYPR-IRCTWIFSOASP--PCEORGLEGYGISKFC 408

Db 315 IFPMINTTVFVNDGENVDLIVEYEAFFPEHQOQWIMYRTFTDKWEDYPKSENEINRIV 374

Qy 409 DHKN-----KPEYIFYAENDDAQFTKMFNTLNIRKQVOLA--NASASQASCSDDGYP 459

Db 375 SELHLRLKGTGGTYTFLVNSDVNAALAFNVYNTKPEILTYDLRVNGLMOCVAAAGFP 434

Qy 460 LPSWTWKCSKSPNCTBEI--PEGVWKKANKRVFGQWVSSSTLNMSEAGKLLVKCCAY 518

Db 435 EPTIDWYFCPCTEQRCASVLPVDVQTLNSSGPPFGKLVQSSIDSSAFKHNGTVECKAY 494

Qy 519 NSMGTSCETIFLNSPGPPFPFIQDN-----ISFYATIGLCLPFIWLVILIC 564

Db 495 NDVGKT--SAYFN-----PAFKGNKEQIHPTLFTPLLIGFVIVAGM---MCIIVMILT 544

Qy 565 HKYKQKRYESQLOMI-QVTGPLDNEYFYVDRDYEDLKWEPFRENLEFGKVLGSAFG 623

Db 545 KYLQKPMYEVQWKVBEING---NNVYIIDPTQLPYDHKEFFPNRLSPFGKTLGAGAFG 601

Qy 624 RVNATAYGISTGVSIOQAVKMLKEKADSEKALMSELKMWTHLGHNDINVLGACT 683

Db 602 KVEATAYGLIKSDAAMTVAVKMLKPSAHLTEREALMSELKVLVSYLGNHMINVLGACT 661

Qy 684 LSGPVYLIFECYCGDILNLYLRSKREKPHRTWTETFEKHNFSSTYPTFOAHNSNMPGSR 743

Db 662 IGGPTLVITEYCCYGDLLNFLRRKRSFICSKQEDHAEALYKYNLLHSHKSSCSDSTNEY 721

Qy 744 VOLHPPDLQD---SGFNNGSIHSEDEIYENQKRLAEEEDLNVLTTFEDLLCPAYQVAK 800

Db 722 MDMPKGVSVVPTKADKRSVRIGSYIERDVTPTAIMEDEALDL---EDLLSFSYQVAK 778

Qy 801 GMEFLFKSCVHRDLAARNVLVTHGKVKIKCDFGLARDILSDSSVVRGNARLPVKWMAP 860

Db 779 GMAFLASKNCIHRDLAARNILLTHGRITKICDFGLARDIKNDSNVVVKGNARLPVKWMAP 838

QY 861 ESLPFIYTKSDVWSYGILLWEIFSLGVNYPGIPVDANFYKLIQSGFMEQPFVATEG 920  
Db 839 ESIFNCVYTFSDVWSYGIFLWELFSLGSPYGMFVDSKFYKMIKGFRLMSPEHAPAE 898  
QY 921 IYFVMSQWAFDSRKRSPFNLTSLFGCOLAEA 953  
Db 899 MIDIMKTCWADPLKRPTEFKQIVOLIEKQISES 931

RESULT 9  
US-09-961-403-4  
; Sequence 4, Application US/09961403  
; Publication NO. US20030077589A1  
; GENERAL INFORMATION:  
; APPLICANT: HE-STUMPP, HOLGER  
; APPLICANT: HAENDLER, BERNARD  
; APPLICANT: KRAETZSCHMAR, JOERN  
; APPLICANT: KREFT, BERTHOLT  
; APPLICANT: WINTERHAGER, ELKE  
; APPLICANT: REGIDOR, PEDRO  
; APPLICANT: SCOTTI, SIMONE  
; TITLE OF INVENTION: METHOD FOR IN VITRO DIAGNOSIS OF ENDOMETRIOSIS  
; FILE REFERENCE: SCH-1789  
; CURRENT APPLICATION NUMBER: US/09/961,403  
; CURRENT FILING DATE: 2001-09-25  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 1088  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-961-403-4

Query Match 21.9%; Score 1152.5; DB 11; Length 1088;  
Best Local Similarity 32.1%; Pred. No. 6.1e-88;  
Matches 314; Conservative 154; Mismatches 320; Indels 189; Gaps 33;

QY 94 VOLATPGDLSCLVWPKHSLGQPHFDLQNRGIVSMALNVTETOAGEY--LLHIQSEEA 151  
Db 39 VQLNSSFSLRCL--FGSEVSNQ--YPMSEESSDVEIRN--EENNSGLFVTVLEVSSASA 92

QY 152 NYTVLFTVNVVTRDTQLYVLRPFYFRKMNQDALLCISEG-----VPEPTV-----EW 197  
Db 93 AHTGLTYTCYNNHQ-----TEENE-----LEGRHIYIYVDPDPAFVPLGMDTY 136

QY 198 VLCSHRES----CK--EEGPAVVRKEEVL-----HELFGTDIRCCARNALGR 240  
Db 137 LVIVEDDSAIIPCRITDPTETVTLHNSGVVPASVDSRQGFNGFTVGPYICEATVRGK 196

QY 241 ECTKLFTIDLNOAPOSTLPQLFL-----KVGEPLMIRKAIHVNHGFLTW-----E 287  
Db 197 ---KQTIIPFNVALKATSELDLEMEALTVYKSGETIVVTC-AVFNNEVDLQWTYPE 252

QY 288 LEDKALEGSEYFEMSTYTNRMIR-----ILLAPVSSVGR---NDTGYYTCSSSK--- 335  
Db 253 VKGKI-----TMLBEIKVPSIKLVYLTVPTEATVKDSDGYECAARQATR 297

QY 336 --HPSQSALVTILEKGF--INATSSOEVEIDPYEKFCSVRFKAYPRIRCTMIFSQASF 391  
Db 298 EVKEMKKVTISVHEKGFIETKFTSQLE--AVNLHEVHKFVFBVRAYPPPRISLWKNLTL 356

QY 392 PCEORGLDGYSIKFCD-----HNNK-----PGEYIFVAENDDAQFTKMTLN 435  
Db 357 -----IENLTETIDVKEIQIRYRSKLKLIRAKEEDSGHYTIVAQNEADVKSITYFELL 410

QY 436 IRKPKOVL-----ANASASQSCSDGYPLPSMTWKKSQKSPNCTEIEIEGVWNKKA 488  
Db 411 TQVPSSILDLVDHSGTGQVTRCTAEGTPLFDIEWMICKD--IKKCNNETS---WTILA 466

QY 489 NRKVPQOWSS-----STLNMSEAGKLLVKKCAVNSMGTSCTIFLNSPG 534  
Db 467 NN-----VSNIIITEIHSRDRSTVGRVTFKAVETIAVRCLAKNLLGAENRELKIVA-- 518

QY 535 PPPFTQDNISFYATTGLCPFFIVLVILCHYKKQKFRIEYSOLOMIQVGTPLDNEYFYVD 594  
Db 519 --PTLRSELTVAAAALVLLVILVILSVIWKQKPRYEIRWRVIESISPDGHEIYVD 576

QY 595 FRDYEYDLKWEFPRENLEFGKVLGSGAFGRVMNATAYGSKTGVSIOQAVKMLKERADSC 654  
Db 577 PMQLPYDSRWEPFRDGLVLRVILGSGAFKGVGTAIGLSRQSPVMKVAVKMLKPTARSS 636

QY 655 EKEALMSLKMTHLGHHDNIYNLLGACLTSGPVVLIFFEYCCYGDLLNLRSLRREKXF-- 711  
Db 637 EKQALMSLKLINTHLGPHLNIYNLLGACTKSGPIYIITEYCFYGDVLYLHNRDPSFLSH 696

QY 712 -----HRTWTEIFKEHNFSSYPTFOAHNSNMP--GSREVLQHP 749  
Db 697 HPEKPKKELDIFGLNPADESTRSYVILSPENNGDYMDMFQADTTQVPMLEKEVSKYS 756

QY 750 LDQLSGFNGNSIHSDEIYEYENQKLABEEEDLNVLTPEDLCLCFAYQVAKGMEFLFKS 809  
Db 757 IQRSYDRPASYYKKKSMULDSEVNKLLSDNSSLTLL---DLLSFTYQVARGMEFLASKN 813

QY 810 CVHRDLAARNVLVTHGKVVKICDFGLARDILDSSYVVRGNARLPVKWMAPESLFEGIYT 869  
Db 814 CVHRDLAARNVLLAOKKIVKICDFGLARDIMHDSNVVSKGSTFLPVKWWAPESIFDNLYT 873

QY 870 IKSDVMSYGILLWEIFSLGVNYPGIPVDANFYKLIQSGFMEQPFYATEGIYFVMSQSW 929  
Db 874 TILSDVMSYGILLWEIFSLGCTPYGPMVMDSTFNKIKSGYRMAKPDHATSEVYEIVVKCW 933

QY 930 APDSRKRPSFPNLTSL 946  
Db 934 NSEPEKRPSFVHLSEIV 950

RESULT 10  
US-09-769-987-2  
; Sequence 2, Application US/09769987  
; Patent No. US20020055129A1  
; GENERAL INFORMATION:  
; APPLICANT: Matsui, Toshimitsu  
; APPLICANT: Aaronson, Stuart A.  
; APPLICANT: Pierce, Jacalyn H.  
; TITLE OF INVENTION: Antibodies for the Alpha Platelet-Derived Growth Factor Receptor  
; FILE REFERENCE: 14014.026602  
; CURRENT APPLICATION NUMBER: US/09/769,987  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 08/460,656  
; PRIOR FILING DATE: 1995-06-02  
; PRIOR APPLICATION NUMBER: US 08/439,095  
; PRIOR FILING DATE: 1995-05-11  
; PRIOR APPLICATION NUMBER: US 07/915,884  
; PRIOR FILING DATE: 1992-07-20  
; PRIOR APPLICATION NUMBER: US 07/308,282  
; PRIOR FILING DATE: 1989-02-09  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 1089  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence; No. US20020055129A1e =  
; OTHER INFORMATION: synthetic construct  
US-09-769-987-2

Query Match 21.9%; Score 1152.5; DB 9; Length 1089;  
Best Local Similarity 32.1%; Pred. No. 6.1e-88;  
Matches 314; Conservative 154; Mismatches 320; Indels 189; Gaps 33;

QY 94 VOLATPGDLSCLVWPKHSLGQPHFDLQNRGIVSMALNVTETOAGEY--LLHIQSEEA 151  
Db 39 VQLNSSFSLRCL--FGSEVSNQ--YPMSEESSDVEIRN--EENNSGLFVTVLEVSSASA 92

QY 152 NYTVLFTVNVVTRDTQLYVLRPFYFRKMNQDALLCISEG-----VPEPTV-----EW 197



Db 814 CVHRDLAARNVLAQKIVKICDFGLARDIMHDSNVSKGSTFLPVKWMAPESIFDNLTYT 873

QY 870 IKSDVMSYGILLWEIFSLGPNYPGIPVDANFYKLIQSGFKMEQPFYATEGYFVVMQSCW 929

Db 874 TLSDVMSYGILLWEIFSLGTPYGMVVDSTFNKIKSGYMAKPDHATSEVYEINVKCW 933

QY 930 AFDSRKRPSPFNLTSL 946

Db 934 NSEPEKRPSPFYHLSEIV 950

RESULT 12

US-09-866-510-2

; Sequence 2, Application US/09866510

; Patent No. US2002011304A1

; GENERAL INFORMATION:

; APPLICANT: KAZLAUSKAS, ANDRIUS

; APPLICANT: IKONO, YASUSHI

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES

; FILE REFERENCE: ERM-104.01

; CURRENT APPLICATION NUMBER: US/09/866,510

; CURRENT FILING DATE: 2001-05-25

; PRIOR APPLICATION NUMBER: 60/250,747

; PRIOR FILING DATE: 2000-12-01

; PRIOR APPLICATION NUMBER: 60/289,103

; PRIOR FILING DATE: 2001-05-07

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: Patent in Ver. 2.1

; SEQ ID NO 2

; LENGTH: 1089

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-866-510-2

Query Match 21.9%; Score 1152.5; DB 10; Length 1089;

Best Local Similarity 32.1%; Pred. No. 6.1e-88;

Matches 314; Conservative 154; Mismatches 320; Indels 189; Gaps 33;

QY 94 VOLATPGDLSCLWVFKHSSIGCOPHPDLQNRGIVSMALNVTETQAGEY--LLHIOSERA 151

Db 39 VQLNSSFSLRC---FGSEVSQW--YPMSEESDVEIRN--EENNSGLFTVLEVSSASA 92

QY 152 NYTVLFTVNRDQLYVLRYPFRKMNQDALLCISEG-----VPEPTY-----EW 197

Db 93 AHTGLTYCYNHTQ-----TEENE-----LEGRHIIYVDPDVAFPVPLGMDTY 136

QY 198: VILCSSHRES---CK---BEGPAVVRKEKVL-----HELCTDIRCCARNALGR 240

Db 137 LVIVEDDDSAIIPCRITDPTVTLHNSGVSPASVDSRQGFNGTFTVGPYICEATVYKG 196

QY 241 ECTKLFITIDNOAPQSTLPOLFL-----KVGRLWIRCKAIHNVHGFGLTW---E 287

Db 197 ---KFTIIPNVVALKATSELDLEMAKTVYKSGETIVTC-AVFNNEVVDLQWTPGE 252

QY 288 LEDKALEGSYPFEMSTYNTMR-----ILLAFVSSYGR---NDTGYTTCSSSK--- 335

Db 253 VKKGGI-----TMLBEIKVPSIKLVTLTVPATVROSGDYECAARQATR 297

QY 336 --HPSQSALVILKQFP--INATSSQBEYIDPYEKFCFVRKAYPRICRTWIFSOASF 391

Db 298 EVKEMKVTIVSHEKGFIEIKPTFSQLE-AVNLHVKHFVVEVRAYPPPRISMLKNLTL 356

QY 392 PCQRQGLDGYSKFCFCD-----HNKK-----PGEYIFYAENDDAQTKMTFLN 435

Db 357 -----IENLTETDVEKIQTIRYSKLLIRAKEDSGHYITVAQNEVAVKSYTFELL 410

QY 436 IRKKPOVL-----ANASASQSCSDGYPLPSWTWKSGDSKSPNCTEPIPEGVWNKKA 488

Db 411 TQVPSSILDLVDDHGGTGGQVTRCTAEGTPLPDIEWMIKCD-IKKCNNETS---WTILA 466

QY 489 NRKVGQWVSS-----STLNMSEAGKGLLVKCCAVNSMGTSCETIFLNSPG 534

Db 467..NN-----VSNIIITEHSRDRSTVEGRVTFPAKVEETIAVRCLAKNLLGAENRELKLV-- 518

QY 535 PPPFTQDNISFVATIGLCLPFIIVLIVLICHYKKQFRIESQLOMIQVTPGLDNEYFYVD 594

Db 519 --PTLRSELTVAAAVLVLIVIIISLIVLWIKQKPYEIRWRVIESIPDGHEIYYVD 576

QY 595 FRDYEYDLKWEPRENLEFGKVLGSGAFGRVNMATAYGISKTGVSIOAVKMLKERADSC 654

Db 577 PMQLPYDSRWEPFRDGLVLRGLSGAFGVGVEGTAYGLSRQPVNMKVAVKMLKPTARRS 636

QY 655 EKEALMSLKMTHLGHHDNIVNLLGACTLSGPFVVLIFEYCCYGDLLNLRSKREKF--- 711

Db 637 EKOALMSLKMTHLGHPLNIVNLLGACTKSGPIYIITEYCFYGDVLYLHKNRDSFLSH 696

QY 712 -----HRTWTEIFKEHNFSSYPTFOAHSNSMP--GSREVQLHPP 749

Db 697 HPEKPKKELDI FGLNPADESTRSYVILSPENNGDYMDMKQADTTQYVPMLEKEVSKYSD 756

QY 750 LDOLSGFNGNSIHSDEIEYENQKLAEEEDLNVLTPEDLCCFAYOVAKGMEFLPKS 809

Db 757 IORSLYDRPASYYKKSMLESEVKNLLSDNSELGLTLL---DLSFTYQVARGMEFLASKN 813

QY 810 CVHRDLAARNVLTGKVVKICDFGLARDIILSDSSVYVRGNARLPVKWMAPESLFEGIYT 869

Db 814 CVHRDLAARNVLAQKIVKICDFGLARDIMHDSNVYKSGSTFLPVKWMAPESIFDNLTYT 873

QY 870 IKSDVMSYGILLWEIFSLGPNYPGIPVDANFYKLIQSGFKMEQPFYATEGYFVVMQSCW 929

Db 874 TLSDVMSYGILLWEIFSLGTPYGMVVDSTFNKIKSGYMAKPDHATSEVYEINVKCW 933

QY 930 AFDSRKRPSPFNLTSL 946

Db 934 NSEPEKRPSPFYHLSEIV 950

RESULT 13

US-09-955-363-36

; Sequence 36, Application US/09955363

; Patent No. US20020173621A1

; GENERAL INFORMATION:

; APPLICANT: Sledziwski Ph.D., Andrzej Z

; Kindvogel Ph.D., Wayne R.

; Bell, Lillian A.

TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS

AND BIOLOGICALLY ACTIVE DIMERIZED POLYPEPTIDE

FUSIONS

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed and Berry

STREET: 6300 Columbia Center

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.24

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/955,363

FILING DATE: 18-Sep-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/634,510

FILING DATE: <Unknown>

APPLICATION NUMBER: US 07/347,291

FILING DATE: 02-MAY-1989

ATTORNEY/AGENT INFORMATION:

NAME: Maki J.D., David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 990008.446C3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206-622-4900



TELEFAX: 206-692-6031  
TELEX: 3723836  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1089 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 36:  
US-09-955-363-36

Query Match 21.9%; Score 1150.5; DB 10; Length 1089;  
Best Local Similarity 32.2%; Pred. No. 9e-88;  
Matches 313; Conservative 154; Mismatches 327; Indels 177; Gaps 32;  
QY 94 VOLATPGDLSCLVFKHSLGCGPHDQLQNRGIVSMALNVTETQAGEY--LLHIQISERA 151  
DB 39 VOLNSSFSLRC---FGESEVSWQ--YPMSEESSDVEIRN--BENNSGLFVTVLEVSSASA 92  
QY 152 NYTVLFTVNRDTQLYVLRPRYFRKMNQDALLCISEG-----VPEPTV-----EW 197  
DB 93 AHTGLTYCYNHTQ-----TEENE-----LEGRHIYIVDPDPAVFPLGMDTY 136  
QY 198 VLCSSHRS-----CK---EKGPAVVRKEEVL-----HELFGTDIRCCARNALGR 240  
DB 137 LVIVEDDDSAIIPCRTPDTPETVTLHNSGVVPASYSRQGFNGTFTVGPYICEATVKGK 196  
QY 241 ECTKLFTIDLNAQPOSTLPQLFL-----KVGEPLWIRCKAIHVNHGFLTW-----E 287  
DB 197 ---KFQTIIPFNVALKATSELDLEMEALKTVYKSGETIVVTC-AVFNNVWDLQWTPGE 252  
QY 288 LEDKALEGSGYFEMSTYNTNRMTIRILLAFVSSVGR---NDTGYTCSSSK-----HPSQ 339  
DB 253 VKGKI-----TILEEIKVPSIKLVTITVPEATVKDSGYECAARQATREKEMK 303  
QY 340 SALVTILEKGF--INATSSQEEYEIDPYEKFCSVRKAYPRIRCTWIPFQASFPCEQRG 397  
DB 304 KVTISVHEKGFIEIKPTFSQLE-AVNLEHVHFVVEVRAVPPRISLWLNKNTL-----356  
QY 398 LEDGYSISKFC-----HKNK-----PGEYIFVAENDDAQFTKMTLNIRKPKQ 441  
DB 357 IENLTITDVEKIOEIRYSKLIIRAKEEDSGHYTIVAQNEHAVKSYTFELLTQVPS 416  
QY 442 VL-----ANASASQACSSDGYPLPSWTKKCDKSPNCTEETIEPEGVWKNKARVFG 494  
DB 417 ILDLVDHGGTGGTGTCTAGTTPDIEIWMICKD-IKKCNETS---WTILAN-----468  
QY 495 QWVSS-----STLNMEAGKGLLVKCCAYNSMTGSCETIFLNSPGPFPIQ 540  
DB 469 --VSNIIITEIHSRDRSTVEGRVTFKAVETIAVRCLAKNLLGAENRELKVA-----522  
QY 541 DNISFYATIGLCPFIVVLVILICHYKQFQRYESQLQMIQVTPGLDNEYFYVDPFRDYEY 600  
DB 523 SELTVAAAVLLVIVILSLVIVIMQKPRYERIRWRVIESISPDGHEIYIVDPQMLPY 582  
QY 601 DLKWEFPRENLEFGVLGSGAFGRVMATAGVISTGYIOAVKMLKEKADSCHEKALM 660  
DB 583 DSRWEFPRDGLVLGRVLGSGAFGVVEGTATGLSRQPVMKVAVKMLKPTARSEKQALM 642  
QY 661 SELKMTHTLGHHDNIVNLLGACTLSGPPVLYLIFEYCCYGDLLNLYRSKEKEF-----711  
DB 643 SELKIMTHLPHLNVNLLGACTSGPIYIITEYCFYGDVLYLHKNRDSFLSHHPEKPK 702  
QY 712 -----HRTWTEIFKEHNFSYPTFOAHNSMPP--GSREVQLHPDLDQLSG 755  
DB 703 KELDIFGLNPADESTRSVILSFENNGDYMOMKQADTTQYVPMLEKEVSKYSDIQRSLY 762  
QY 756 FNGNSIHSDEIEYENQKRLAEEEDLNLTFEDLLCFAYQVAKGMFLEFKSCVHRDL 815  
DB 763 DRPASYKKKMLDSEVKNLLSDDDSEGLTLL---DLSFTYQVARGMEFLASKNCVHRDL 819  
QY 816 AARNVLTHGVKVKICDFGLARDILSDSSVYVGRNARLPVKWMAPESEIFDNLYTTLSDVM 875

DB 820 AARNVLLAQKIVKICDFGLARDIMHDSNVVSKGSTFLPVKWMAPESIFDNLYTTLSDVM 879  
QY 876 SYGILLWEIFSLGVNPNYRGIPVDANFYKLIQSGFKMEQPFYATEGTYFVQMSWAFDSRK 935  
DB 880 SYGILLWEIFSLGCTYPCGMVMDSTFYNKIKSGYRMAKPDHATSEVYEIMVKWNSPEK 939  
QY 936 RPSFNPNTSFL 946  
DB 940 RPSFYHLSEIV 950  
RESULT 14  
US-09-866-510-10  
; Sequence 10, Application US/09866510  
; Patent No. US20020111304A1  
; GENERAL INFORMATION:  
; APPLICANT: KAZLAUSKAS, ANDRIUS  
; APPLICANT: IKUNO, YASUSHI  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES  
; FILE REFERENCE: ERM-104.01  
; CURRENT APPLICATION NUMBER: US/09/866,510  
; CURRENT FILING DATE: 2001-05-25  
; PRIOR APPLICATION NUMBER: 60/250,747  
; PRIOR FILING DATE: 2000-12-01  
; PRIOR APPLICATION NUMBER: 60/289,103  
; PRIOR FILING DATE: 2001-05-07  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 1089  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-866-510-10  
Query Match 21.8%; Score 1149.5; DB 10; Length 1089;  
Best Local Similarity 32.0%; Pred. No. 1.1e-87;  
Matches 313; Conservative 155; Mismatches 320; Indels 189; Gaps 33;  
QY 94 VOLATPGDLSCLVFKHSLGCGPHDQLQNRGIVSMALNVTETQAGEY--LLHIQISERA 151  
DB 39 VOLNSSFSLRC---FGESEVSWQ--YPMSEESSDVEIRN--BENNSGLFVTVLEVSSASA 92  
QY 152 NYTVLFTVNRDTQLYVLRPRYFRKMNQDALLCISEG-----VPEPTV-----EW 197  
DB 93 AHTGLTYCYNHTQ-----TEENE-----LEGRHIYIVDPDPAVFPLGMDTY 136  
QY 198 VLCSSHRS-----CK---EKGPAVVRKEEVL-----HELFGTDIRCCARNALGR 240  
DB 137 LVIVEDDDSAIIPCRTPDTPETVTLHNSGVVPASYSRQGFNGTFTVGPYICEATVKGK 196  
QY 241 ECTKLFTIDLNAQPOSTLPQLFL-----KVGEPLWIRCKAIHVNHGFLTW-----E 287  
DB 197 ---KFQTIIPFNVALKATSELDLEMEALKTVYKSGETIVVTC-AVFNNVWDLQWTPGE 252  
QY 288 LEDKALEGSGYFEMSTYNTNRMTIR-----ILLAFVSSVGR---NDTGYTCSSSK---335  
DB 253 VKGKI-----TILEEIKVPSIKLVTITVPEATVKDSGYECAARQATR 297  
QY 336 --HPSQASALVILEKGF--INATSSQEEYEIDPYEKFCSVRKAYPRIRCTWIPFQASFP 391  
DB 298 EVKEMKVTISVHEKGFIEIKPTFSQLE-AVNLEHVHFVVEVRAVPPRISLWLNKNTL 356  
QY 392 PCEQRGLDGYISKFC-----HKNK-----PGEYIFVAENDDAQFTKMTLN 435  
DB 357 -----IENLTITDVEKIOEIRYSKLIIRAKEEDSGHYTIVAQNEHAVKSYTFELL 410  
QY 436 IRKPKQVL-----ANASASQACSSDGYPLPSWTKKCDKSPNCTEETIEPEGVWKNKA 488  
DB 411 TQVPSILLVDHGGTGGTGTCTAGTTPDIEIWMICKD-IKKCNETS---WTILA 466  
QY 489 NRKYPGQWVSS-----STLNMEAGKGLLVKCCAYNSMTGSCETIFLNSPG 534  
DB 467 NN-----VSNIIITEIHSRDRSTVEGRVTFKAVETIAVRCLAKNLLGAENRELKVA--518

```
QY 535 PFFIQDNISFYATIGLCLPFIIVLIVLICHYKQFYESQLOMIQVGTPLDNEYFYVD 594
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 519 --PTLRSELVAAAVLLVLIISLIVLVVIMQKPRYEIRWRVIBSISPDGHEIYVD 576
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 595 PRDYEYDLKWEFFPRENLLEFGVLGSGAGRVVGNATAGISKTGYSIOVAVKMLKEKADSC 654
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 577 PMQLPYDSRWEPFRDGLVLRVLSGAGFGKVEGTAYGLSRSPQPMKVAVKMLKPTARSS 636
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 655 EKEALMSELKMWTHLGHHDNIWNLLGACTLSGPPVLIPEYCYGDLNLYLRSKREKF--- 711
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 637 EKOALMSELKMWTHLGHPLNIVNLLGACTKSGPIIITEYCFYGDVNLVHLKRNDSFLSH 696
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 712 -----HRTWTEIFKEHNFSSYPTFOAHNSMMP--GSRVQLHPP 749
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 697 HPEKPKKELDIFGLNPADESTRSVILSFENNMDYMDMKQADTTQYVPMLEKREKYSYD 756
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 750 LDQLSGFNGNSIHSEDEIEYENOKRLAEEBEDNLVTFEDLLCFAYOVAKGMEFLPKS 809
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 757 IORSYDRPASYYKKKSMLESDSEKLLSDONSEGLTLL---DLLSTYQVARGMEFLASKN 813
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 810 CVHRDLAARNVLTGHKVKVVICDFGLARDILSDGSYVVRGNARLPVKWMAPESLFEGIYT 869
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 814 CVHRDLAARNVLLAOGKIVKICDFGLARDIMHDSNYVSKGSTFLPMKMAPESLFDNLYT 873
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 870 IKSDVWSYGILLWEIFSLGVNPNYPGIPVDANFYKLIQSGFKMQPFYATEGIYFVMSQSW 929
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 874 TLDVWSYGILLWEIFSLGGTFYFGMMVYDSTFYNKIKSGYRMAKPDHATSEVYIEMVKCW 933
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 930 AFDSRKRPSPFNLTSL 946
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 934 NSEPEKRPSPFYLSEIV 950
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
US-09-866-510-4
; Sequence 4, Application US/09866510
; Patent No. US20020111304A1
; GENERAL INFORMATION:
; APPLICANT: KAZLAUSKAS, ANDRIUS
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
; FILE REFERENCE: ERM-104.01
; CURRENT APPLICATION NUMBER: US/09/866.510
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/250,747
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/289,103
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1.
; SEQ ID NO 4
; LENGTH: 1089
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-866-510-4

Query Match 21.8%; Score 1148.5; DB 10; Length 1089;
Best Local Similarity 32.0%; Pred. No. 1.3e-87;
Matches 313; Conservative 155; Mismatches 320; Indels 189; Gaps 33;

QY 94 VOLATPGDLSCLWVFKHSLGCGPHFDLQNRGIVSMALNVTTQAGEY--LLHIQSER 151
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 39 VOLANSFSLRC---PGESEVSWQ--YPMSEBSSDVEIRN--EENNSGLFVTVLEVSSASA 92
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 152 NYTVLFTVNVNRDTQLYVLRPRYFRKQENODALLCTISEG-----VPEPTV-----EW 197
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 93 AHTGLYTCYNNHTQ-----TEENE-----LEGRHIYIYVDPDPVAFVPLGMDTY 136
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 198 VLCSSHRES---CK---EEGPAVVRKEEKL-----HELFGTDIRCCARNALGR 240
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 137 LVIVEDDDSAIIPCKRTTDPETVTLHNSEGVVPASYSRQGFNGFTTVGPGYICEATVRGK 196
```

```
QY 241 ECTKLTIDLNQAPQSTLPQLFL-----KVEBPLWIRCKAIHVNHGFLTW-----E 287
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 197 ---KFQTFIPNVVALKATSELDLEMEALKTVYKSETIVVTC-AVFENNEVDLQMTYPE 252
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 288 LEDKALBEGSYFEMSTYSTNRTMIR-----ILLAFVSSVGR---NDTGYTCSSSK--- 335
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 253 VKGKI-----TWLEEKVPSIKLVYTLTVPETVVKDSGDYECAARQATR 297
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 336 --HPSQSALVTILEKGF--INATSSQBEYEIDPEKFCFSVRFKAYPRICRTWIIFSQASF 391
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 298 EVKEMKVTISVHEKGFIEIKPTPSQLE-AVNLHEVKHFVVEVRAYPPPRISLWKNLTL 356
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 392 PCEORGLEDGYSISKPCD-----HKNK-----PGEYIFYAENDDAQFTQWFTLN 435
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 357 -----IENLTEITTEVKEIOEIRSKLUIRAKEDSGHYTTVAQEDAVKSTFELL 410
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 436 IRKKPQVL-----ANASASQACSSDGYPLPSWTWKCKSDKSPNCTEETEEIPEGVWNKA 488
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 411 TQVPSSILDLVDDHGGTGGTGTAEGLTLPDIEWMICKD-IKCKNNETS---WTLA 466
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 489 NRKVFQGWVSS-----STLNMSEAKGGLLVKCCAYNSMGTSCETIFLNSPG 534
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 467 NN-----VSNIIIEHSDRSTVEGRVTFAKVEETIAVRCLAKNLLGAENRELKVA-- 518
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 535 PFFIQDNISFYATIGLCLPFIIVLIVLICHYKQFYESQLOMIQVGTPLDNEYFYVD 594
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 519 --PTLRSELVAAAVLLVLIISLIVLVVIMQKPRYEIRWRVIBSISPDGHEIYVD 576
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 595 PRDYEYDLKWEFFPRENLLEFGVLGSGAGRVVGNATAGISKTGYSIOVAVKMLKEKADSC 654
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 577 PMQLPYDSRWKFPDRDGLVLRVLSGAGFGKVEGTAYGLSRSPQPMKVAVKMLKPTARSS 636
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 655 EKEALMSELKMWTHLGHHDNIWNLLGACTLSGPPVLIPEYCYGDLNLYLRSKREKF--- 711
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 637 EKOALMSELKMWTHLGHPLNIVNLLGACTKSGPIIITEYCFYGDVNLVHLKRNDSFLSH 696
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 712 -----HRTWTEIFKEHNFSSYPTFOAHNSMMP--GSRVQLHPP 749
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 697 HPEKPKKELDIFGLNPADESTRSVILSFENNMDYMDMKQADTTQYVPMLEKREKYSYD 756
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 750 LDQLSGFNGNSIHSEDEIEYENOKRLAEEBEDNLVTFEDLLCFAYOVAKGMEFLPKS 809
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 757 IORSYDRPASYYKKKSMLESDSEKLLSDONSEGLTLL---DLLSTYQVARGMEFLASKN 813
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 810 CVHRDLAARNVLTGHKVKVVICDFGLARDILSDGSYVVRGNARLPVKWMAPESLFEGIYT 869
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 814 CVHRDLAARNVLLAOGKIVKICDFGLARDIMHDSNYVSKGSTFLPMKMAPESLFDNLYT 873
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 870 IKSDVWSYGILLWEIFSLGVNPNYPGIPVDANFYKLIQSGFKMQPFYATEGIYFVMSQSW 929
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 874 TLDVWSYGILLWEIFSLGGTFYFGMMVYDSTFYNKIKSGYRMAKPDHATSEVYIEMVKCW 933
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 930 AFDSRKRPSPFNLTSL 946
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 934 NSEPEKRPSPFYLSEIV 950
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Search completed: August 26, 2003, 07:41:07
Job time : 49.4771 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2003, 07:21:59 ; Search time 21.4892 Seconds  
(without alignments)  
1953.187 Million cell updates/sec

Title: US-09-919-408A-2

Perfect score: 5264

Sequence: 1 MRALAQRSDRLLLLVLSV.....RGGLRAQSPQKVIHRERS 992

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*\*

- 1: /cgn2\_6/ptodata/1/iaa/5A.COMB.pap:\*\*
- 2: /cgn2\_6/ptodata/1/iaa/5B.COMB.pap:\*\*
- 3: /cgn2\_6/ptodata/1/iaa/6A.COMB.pap:\*\*
- 4: /cgn2\_6/ptodata/1/iaa/6B.COMB.pap:\*\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS.COMB.pap:\*\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pap:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5264	100.0	992	1	US-07-813-593-2
2	5264	100.0	992	1	US-07-977-451-2
3	5264	100.0	992	1	US-07-946-507-2
4	5264	100.0	992	1	US-08-252-517-2
5	5264	100.0	992	1	US-07-906-397A-2
6	5264	100.0	992	1	US-08-601-891-2
7	5264	100.0	992	2	US-09-021-324-2
8	5264	100.0	992	5	PCT-US92-02750-2
9	5264	100.0	992	5	PCT-US92-05401-2
10	5264	100.0	992	5	PCT-US92-09893-2
11	5102	96.9	1000	1	US-08-222-299-2
12	5102	96.9	1000	1	US-08-434-878-2
13	5102	96.9	1000	1	PCT-US95-03718-2
14	4436.5	84.3	993	1	US-08-222-299-4
15	4436.5	84.3	993	2	US-08-434-878-4
16	4436.5	84.3	993	5	PCT-US95-03718-4
17	4436.5	84.2	1160	5	PCT-US92-05401-4
18	4429.5	84.1	993	1	US-07-977-451-4
19	4429.5	84.1	993	1	US-08-252-517-4
20	4429.5	84.1	993	1	US-07-906-397A-4
21	4429.5	84.1	993	1	US-08-601-891-4
22	4429.5	84.1	993	2	US-09-021-324-4
23	4429.5	84.1	993	5	PCT-US92-09893-4
24	4408.5	83.7	993	1	US-08-183-211-2
25	4408.5	83.7	993	5	PCT-US95-00176A-2
26	2240.5	42.6	481	3	US-07-912-122-4
27	2240.5	42.6	481	5	PCT-US93-06404-4

Query Match 100.0%; Score 5264; DB 1; Length 992;

US-07-813-593-2  
; Sequence 2, Application US/07813593  
; Patent No. 5185438  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMCLONE SYSTEMS INCORPORATED  
; STREET: 180 VARICK STREET  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/813,593  
; FILING DATE: 19920415  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/793,065  
; FILING DATE: 15-NOV-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/728,913  
; FILING DATE: 28-JUN-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/679,666  
; FILING DATE: 02-APR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Feit, Irving N.  
; REGISTRATION NUMBER: 28,601  
; REFERENCE/DOCKET NUMBER: LEW-3-PPP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-645-1405  
; TELEFAX: 212-645-2054  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 992 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-813-593-2

## ALIGNMENTS

### RESULT 1

US-07-813-593-2  
; Sequence 2, Application US/07813593  
; Patent No. 5185438  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMCLONE SYSTEMS INCORPORATED  
; STREET: 180 VARICK STREET  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/813,593  
; FILING DATE: 19920415  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/793,065  
; FILING DATE: 15-NOV-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/728,913  
; FILING DATE: 28-JUN-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/679,666  
; FILING DATE: 02-APR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Feit, Irving N.  
; REGISTRATION NUMBER: 28,601  
; REFERENCE/DOCKET NUMBER: LEW-3-PPP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-645-1405  
; TELEFAX: 212-645-2054  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 992 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-813-593-2

Query Match 100.0%; Score 5264; DB 1; Length 992;

Best Local Similarity 100.0%; Pred. No. 0;  
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALQSRDRRLLLVLSVMILETVTNQDLPVVKVLSHENNGSSAGKPSYSYRMVGRS 60  
Db 1 MRALQSRDRRLLLVLSVMILETVTNQDLPVVKVLSHENNGSSAGKPSYSYRMVGRS 60

QY 61 PEDLOCTPRQSEGTYVEAATVEAEGSGITLQVQLATPGDLSCLVWFKHSSIGCQPHFD 120  
Db 61 PEDLOCTPRQSEGTYVEAATVEAEGSGITLQVQLATPGDLSCLVWFKHSSIGCQPHFD 120

QY 121 LQNRGIVSMALNVETQAGEYLLHIQSERANYVLTFTVNRDQTOYLYVLRPYFRKMNQ 180  
Db 121 LQNRGIVSMALNVETQAGEYLLHIQSERANYVLTFTVNRDQTOYLYVLRPYFRKMNQ 180

QY 181 DALLCISEGVPEPTVEWVLCSSHRESCKEEGPAVVRKEKVLHELFGTDIRCARNALGR 240  
Db 181 DALLCISEGVPEPTVEWVLCSSHRESCKEEGPAVVRKEKVLHELFGTDIRCARNALGR 240

QY 241 ECTKLTIDLNQAPOSTLPQLFLKVGEPWIRCKAIHVNHGFGLTWELEDKALEEGSYFE 300  
Db 241 ECTKLTIDLNQAPOSTLPQLFLKVGEPWIRCKAIHVNHGFGLTWELEDKALEEGSYFE 300

QY 301 MSTYSTNRTMIRILLAFVSSVGRNDGYTTCSSKHPQSALVTILEKGFINATSQEY 360  
Db 301 MSTYSTNRTMIRILLAFVSSVGRNDGYTTCSSKHPQSALVTILEKGFINATSQEY 360

QY 361 EIDPYEKCFVSFRKAYPIRITCTWFSQAFCEQRLGLEDGYISIKFCDHKPKGEYIEY 420  
Db 361 EIDPYEKCFVSFRKAYPIRITCTWFSQAFCEQRLGLEDGYISIKFCDHKPKGEYIEY 420

QY 421 AENDDAQFTKMTLIRKPKQVLNANASQASCSDDGYPLPGMTWKCKDKSPNCTEETIP 480  
Db 421 AENDDAQFTKMTLIRKPKQVLNANASQASCSDDGYPLPGMTWKCKDKSPNCTEETIP 480

QY 481 EGVNKKANRKYFGQWSSSTLNMEAGKGLVKKCCAYNSMGTSCETIFLNSPGPPFTQ 540  
Db 481 EGVNKKANRKYFGQWSSSTLNMEAGKGLVKKCCAYNSMGTSCETIFLNSPGPPFTQ 540

QY 541 DNISFYATIGLCPLFIVLVILCHYKKQFYESQLOMIQVTPGLDNEYFYVDFRDY 600  
Db 541 DNISFYATIGLCPLFIVLVILCHYKKQFYESQLOMIQVTPGLDNEYFYVDFRDY 600

QY 601 DLKWEFFPRENLBFGLSGAFGRVWNAAYGISKTVGSIQVAVKMLKEKADSCKEALM 660  
Db 601 DLKWEFFPRENLBFGLSGAFGRVWNAAYGISKTVGSIQVAVKMLKEKADSCKEALM 660

QY 661 SELKMTWHLGHDNIIVNLLGACTLGGPVYLI FEYCCYGDLLNLYRSKREKFKHRTWTEIFK 720  
Db 661 SELKMTWHLGHDNIIVNLLGACTLGGPVYLI FEYCCYGDLLNLYRSKREKFKHRTWTEIFK 720

QY 721 EHNFSYPTFOAHSSNMPGSEVQLHPDLQSGFNNGSIHSEDEIEYENQKLAEBEE 780  
Db 721 EHNFSYPTFOAHSSNMPGSEVQLHPDLQSGFNNGSIHSEDEIEYENQKLAEBEE 780

QY 781 EDNLVLTEDLLCFAYQVAKGMEFLFKSCVRDLAARNVLVTHGKVKICDFGLARDIL 840  
Db 781 EDNLVLTEDLLCFAYQVAKGMEFLFKSCVRDLAARNVLVTHGKVKICDFGLARDIL 840

QY 841 SDSSYVVRGNARLPVKWMAPELFGIYTIKSDVMSYGLLWEIFSLGVNYPYGPVVDAN 900  
Db 841 SDSSYVVRGNARLPVKWMAPELFGIYTIKSDVMSYGLLWEIFSLGVNYPYGPVVDAN 900

QY 901 FYKLIQSGFKMBQPYATYEGYFVWQSCWAFPSRKRPSFNLTSFLGCOLAABEACIRT 960  
Db 901 FYKLIQSGFKMBQPYATYEGYFVWQSCWAFPSRKRPSFNLTSFLGCOLAABEACIRT 960

QY 961 SIHLPKQAAPQORGGLRAQSPORQVKIHERS 992  
Db 961 SIHLPKQAAPQORGGLRAQSPORQVKIHERS 992

RESULT 2

US-07-977-451-2  
; Sequence 2, Application US/07977451  
; Patent No. 5270458  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSES: ImClone Systems Incorporated  
; STREET: 180 Varick Street  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/977,451  
; FILING DATE: 19921119  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US UNASSIGNED  
; FILING DATE: 12-NOV-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/906,397  
; FILING DATE: 26-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US PCT/US92/05401  
; FILING DATE: 26-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: TW 81102961  
; FILING DATE: 15-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US PCT/US92/02750  
; FILING DATE: 02-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/813,593  
; FILING DATE: 24-DEC-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/793,065  
; FILING DATE: 15-NOV-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/728,913  
; FILING DATE: 28-JUN-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/679,666  
; FILING DATE: 02-APR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Feit, Irving N.  
; REGISTRATION NUMBER: 28,601  
; REFERENCE/DOCKET NUMBER: LEM-3-7P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-645-1405  
; TELEFAX: 212-645-2054  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 992 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-977-451-2

Query Match 100.0%; Score 5264; DB 1; Length 992;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALQSRDRRLLLVLSVMILETVTNQDLPVVKVLSHENNGSSAGKPSYSYRMVGRS 60  
Db 1 MRALQSRDRRLLLVLSVMILETVTNQDLPVVKVLSHENNGSSAGKPSYSYRMVGRS 60

Qy 61 PEDLOCTPRROSEGTVEAATVEAESSGSIITLOVLATPGDLSCLVFVFKHSSLGCPHFD 120  
Db 61 PEDLOCTPRROSEGTVEAATVEAESSGSIITLOVLATPGDLSCLVFVFKHSSLGCPHFD 120  
Qy 121 LONRGIVSMALINVTETQAGEYLLHIQSERANYTVLFTVNVVDTQLYVLRPPYFRMENO 180  
Db 121 LONRGIVSMALINVTETQAGEYLLHIQSERANYTVLFTVNVVDTQLYVLRPPYFRMENO 180  
Qy 181 DALLCISGEVPEPTVEWVLCSSHRESCKEGBPAVVRKEEVLHFGTDIRCCARNALGR 240  
Db 181 DALLCISGEVPEPTVEWVLCSSHRESCKEGBPAVVRKEEVLHFGTDIRCCARNALGR 240  
Qy 241 ECTKLFITDLNQAPOSTLPOLFLKVGEPWIRCKAIHVNHGFGLTWELEDKALEGSYFE 300  
Db 241 ECTKLFITDLNQAPOSTLPOLFLKVGEPWIRCKAIHVNHGFGLTWELEDKALEGSYFE 300  
Qy 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTTCSSSKHPSQSALVTILEKGFINATSSQBEY 360  
Db 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTTCSSSKHPSQSALVTILEKGFINATSSQBEY 360  
Qy 361 BIDPYEKECFVSFRKAYPRIRCTWIFSQASFPCEORGLDGYISIKFCDHKNKPGEYIFY 420  
Db 361 BIDPYEKECFVSFRKAYPRIRCTWIFSQASFPCEORGLDGYISIKFCDHKNKPGEYIFY 420  
Qy 421 AENDDAQFTKFTNIRKPPQVLNANASASQASCSGDPYPLPSWTWKKCSDKSPNCTEIP 480  
Db 421 AENDDAQFTKFTNIRKPPQVLNANASASQASCSGDPYPLPSWTWKKCSDKSPNCTEIP 480  
Qy 481 EGVNKKANRVFGOWVSSSTLNSEAGKLLVKCAVNSMTGTCETIFLNSPGPPFIQ 540  
Db 481 EGVNKKANRVFGOWVSSSTLNSEAGKLLVKCAVNSMTGTCETIFLNSPGPPFIQ 540  
Qy 541 DNISFYATIGCLPPIVVLIVLICHKYYKQFRYESQLOMIQVTPGLDNEYFYVDFRDY 600  
Db 541 DNISFYATIGCLPPIVVLIVLICHKYYKQFRYESQLOMIQVTPGLDNEYFYVDFRDY 600  
Qy 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISKTGVSIOVAVQMLKEKADSCKEALM 660  
Db 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISKTGVSIOVAVQMLKEKADSCKEALM 660  
Qy 661 SELKQWTHLGHNDINVLGACTLSGPVYLIFEYCCYGDLLNLYRSKEKPHRTWTEIFK 720  
Db 661 SELKQWTHLGHNDINVLGACTLSGPVYLIFEYCCYGDLLNLYRSKEKPHRTWTEIFK 720  
Qy 721 EHNFSYPTFOAHNSNMPGSRVQLPPLDQLSGFNGNSIHSEDEIYENQKRLAEBEE 780  
Db 721 EHNFSYPTFOAHNSNMPGSRVQLPPLDQLSGFNGNSIHSEDEIYENQKRLAEBEE 780  
Qy 781 EDNLVTFEDLLCFAYQVAKWMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840  
Db 781 EDNLVTFEDLLCFAYQVAKWMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840  
Qy 841 SDSYVVRGNARLPVKWMAPELSEGIYTIKSDVMSYIGILLWEIFSLGWNYPGIPVDAN 900  
Db 841 SDSYVVRGNARLPVKWMAPELSEGIYTIKSDVMSYIGILLWEIFSLGWNYPGIPVDAN 900  
Qy 901 FYKLIQSGFKMEQPFYATEGIFYVWQSCWAFDSRKRPSFNLTSFLGCOLAEBEACIRT 960  
Db 901 FYKLIQSGFKMEQPFYATEGIFYVWQSCWAFDSRKRPSFNLTSFLGCOLAEBEACIRT 960  
Qy 961 SIHLPKQAAPOQRGGLRAQSPQRQVKIHRERS 992  
Db 961 SIHLPKQAAPOQRGGLRAQSPQRQVKIHRERS 992

RESULT 3  
US-07-946-507-2  
; Sequence 2, Application US/07946507  
; Patent No. 5283354  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL

; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMCLONE SYSTEMS INCORPORATED  
; STREET: 180 VARICK STREET  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/946,507  
; FILING DATE: 19920917  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/813,593  
; FILING DATE: 24-DEC-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/793,065  
; FILING DATE: 15-NOV-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/728,913  
; FILING DATE: 28-JUN-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/679,666  
; FILING DATE: 02-APR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Feit, Irving N.  
; REGISTRATION NUMBER: 28,601  
; REFERENCE/DOCKET NUMBER: LEM-3-PPP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-645-1405  
; TELEFAX: 212-645-2054  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 992 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-07-946-507-2

Query Match 100.0%; Score 5264; DB 1; Length 992;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MRALQASDRRLLLVLSVMILETVTNQDLPVVKCVLISHENNGSSAGKPSYSYRWVRS 60  
Db 1 MRALQASDRRLLLVLSVMILETVTNQDLPVVKCVLISHENNGSSAGKPSYSYRWVRS 60  
Qy 61 PEDLOCTPRROSEGTVEAATVEAESSGSIITLOVLATPGDLSCLVFVFKHSSLGCPHFD 120  
Db 61 PEDLOCTPRROSEGTVEAATVEAESSGSIITLOVLATPGDLSCLVFVFKHSSLGCPHFD 120  
Qy 121 LONRGIVSMALINVTETQAGEYLLHIQSERANYTVLFTVNVVDTQLYVLRPPYFRMENO 180  
Db 121 LONRGIVSMALINVTETQAGEYLLHIQSERANYTVLFTVNVVDTQLYVLRPPYFRMENO 180  
Qy 181 DALLCISGEVPEPTVEWVLCSSHRESCKEGBPAVVRKEEVLHFGTDIRCCARNALGR 240  
Db 181 DALLCISGEVPEPTVEWVLCSSHRESCKEGBPAVVRKEEVLHFGTDIRCCARNALGR 240  
Qy 241 ECTKLFITDLNQAPOSTLPOLFLKVGEPWIRCKAIHVNHGFGLTWELEDKALEGSYFE 300  
Db 241 ECTKLFITDLNQAPOSTLPOLFLKVGEPWIRCKAIHVNHGFGLTWELEDKALEGSYFE 300  
Qy 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTTCSSSKHPSQSALVTILEKGFINATSSQBEY 360  
Db 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTTCSSSKHPSQSALVTILEKGFINATSSQBEY 360

```
QY 361 EIDPYEKFCFSVRFKAYPRIRCTWIFSOASFPCEQGLGDSYISKFCCHKPKGEYIFY 420
Db 361 EIDPYEKFCFSVRFKAYPRIRCTWIFSOASFPCEQGLGDSYISKFCCHKPKGEYIFY 420
QY 421 AENDDAQFTKMTLIRKPKQVLANASQASCSGSDGYPLPSWTWKKCSKSPNCTEIP 480
Db 421 AENDDAQFTKMTLIRKPKQVLANASQASCSGSDGYPLPSWTWKKCSKSPNCTEIP 480
QY 481 EGVNKKANRKFVGOVSSSTLNMEAGKLLVKCCAYNSMGTSCETIFLNSPGFPFIQ 540
Db 481 EGVNKKANRKFVGOVSSSTLNMEAGKLLVKCCAYNSMGTSCETIFLNSPGFPFIQ 540
QY 541 DNISFYATIGLCLPFIIVLIVLICHYKQFYESOLQMIQVTPLDNEYFYVDRDYEY 600
Db 541 DNISFYATIGLCLPFIIVLIVLICHYKQFYESOLQMIQVTPLDNEYFYVDRDYEY 600
QY 601 DLKWEFPRENLEFGKVLGSGAFGRVYNATAYGISKTVGSIQVAVKMLKEKADSCKEALM 660
Db 601 DLKWEFPRENLEFGKVLGSGAFGRVYNATAYGISKTVGSIQVAVKMLKEKADSCKEALM 660
QY 661 SELKMTLHGHNDINVLGACTLSPVYLIFEYCCYGDLLNLYRSKREKPHRTWTETPK 720
Db 661 SELKMTLHGHNDINVLGACTLSPVYLIFEYCCYGDLLNLYRSKREKPHRTWTETPK 720
QY 721 EHNFSYPTFOAHSNSMPGSRVQLHPDLQSLGFGNGSIHSEDEIEYENOKRLAEBEE 780
Db 721 EHNFSYPTFOAHSNSMPGSRVQLHPDLQSLGFGNGSIHSEDEIEYENOKRLAEBEE 780
QY 781 EDNLVLTEDLLCFAYQVAKMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
Db 781 EDNLVLTEDLLCFAYQVAKMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
QY 841 SDSSVYVGNARLPVKWMAPELSFEGIYTIKSDVMSYGILLWEIFSLGWNVPPIPVDA 900
Db 841 SDSSVYVGNARLPVKWMAPELSFEGIYTIKSDVMSYGILLWEIFSLGWNVPPIPVDA 900
QY 901 FYKLQSGFKMEQPFYATIGIYFVMSQWAFDSRKPSPNLTSLFGCOLABAEACIRT 960
Db 901 FYKLQSGFKMEQPFYATIGIYFVMSQWAFDSRKPSPNLTSLFGCOLABAEACIRT 960
QY 961 SIHLPKQAAPQORGGLRAQSPORQVKIHRERS 992
Db 961 SIHLPKQAAPQORGGLRAQSPORQVKIHRERS 992

RESULT 4
US-08-252-517-2
; Sequence 2, Application US/08252517
; Patent No. 5548065
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/252,517
; FILING DATE: 31-OCT-1994
; CLASSIFICATION: 530
; -PRIOR APPLICATION NUMBER: US 07/977,451
; FILING DATE: 19-NOV-1992
```

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-7P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 992 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-252-517-2

Query Match 100.0%; Score 5264; DB 1; Length 992;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEALAQRSDRRLLLVLSVMILETTVTNQDLPEIKVLISHENHGSSACKPSYRMVRGS 60
Db 1 MEALAQRSDRRLLLVLSVMILETTVTNQDLPEIKVLISHENHGSSACKPSYRMVRGS 60
QY 61 PEDLOCTPRQSEGTVEAATVEVAESGSIITLQVQLATPGDLSCLVFKHSLGCPHF 120
Db 61 PEDLOCTPRQSEGTVEAATVEVAESGSIITLQVQLATPGDLSCLVFKHSLGCPHF 120
QY 121 LQNRGIVSMAILNVTTQAGEYLLHIQSERANYTLFTVNVVDTOLYVLRPPYFRKWEQ 180
Db 121 LQNRGIVSMAILNVTTQAGEYLLHIQSERANYTLFTVNVVDTOLYVLRPPYFRKWEQ 180
QY 181 DALLCISEGVPEPTVEWVLCSSHRSCKEEGPAVVRKEEKVLHELFGTDIRCCARNALGR 240
Db 181 DALLCISEGVPEPTVEWVLCSSHRSCKEEGPAVVRKEEKVLHELFGTDIRCCARNALGR 240
QY 241 ECTKLFTIDLNOAPOSTLPQLFKVGEPLWIRKAIHNVHNGFGLTWELDKALEEGSYFE 300
Db 241 ECTKLFTIDLNOAPOSTLPQLFKVGEPLWIRKAIHNVHNGFGLTWELDKALEEGSYFE 300
QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCCSSSKHPSOSALVTILEKGFINATSSQEEY 360
Db 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCCSSSKHPSOSALVTILEKGFINATSSQEEY 360
QY 361 EIDPYEKFCFSVRFKAYPRIRCTWIFSOASFPCEQGLGDSYISKFCCHKPKGEYIFY 420
Db 361 EIDPYEKFCFSVRFKAYPRIRCTWIFSOASFPCEQGLGDSYISKFCCHKPKGEYIFY 420
QY 421 AENDDAQFTKMTLIRKPKQVLANASQASCSGSDGYPLPSWTWKKCSKSPNCTEIP 480
Db 421 AENDDAQFTKMTLIRKPKQVLANASQASCSGSDGYPLPSWTWKKCSKSPNCTEIP 480
```

Db 421 AENDDAQFTKMTFLNIRKKPQVLNANASQASCSGYPPLPSWTWKCKSDKSPNCTEIP 480  
Qy 481 EGVNKKANRKYFGQWSSSTLNSEAGKLLVKKCCAYNSMGTSCTETFLNSGPPPTQ 540  
Db 481 EGVNKKANRKYFGQWSSSTLNSEAGKLLVKKCCAYNSMGTSCTETFLNSGPPPTQ 540  
Qy 541 DNISFYATIGLCPPIVVLIVLICHYKQKQFYESQLOMIQVTGPDNDNEYFYVDFRDY 600  
Db 541 DNISFYATIGLCPPIVVLIVLICHYKQKQFYESQLOMIQVTGPDNDNEYFYVDFRDY 600  
Qy 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISTGVSIOVAVKMLKEKADSCKEALM 660  
Db 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISTGVSIOVAVKMLKEKADSCKEALM 660  
Qy 661 SELKQWTHLGHNDINVLGACTLSPVYLIIEYCCYGDLLNLSKREKFRHTWTEIPK 720  
Db 661 SELKQWTHLGHNDINVLGACTLSPVYLIIEYCCYGDLLNLSKREKFRHTWTEIPK 720  
Qy 721 EHNFSYPTFOAHNSNMPGSRVQLHPPDLQSGFNGNSIHSDEIEYENOKRLAEERE 780  
Db 721 EHNFSYPTFOAHNSNMPGSRVQLHPPDLQSGFNGNSIHSDEIEYENOKRLAEERE 780  
Qy 781 EDNLVLTPEDDLCPAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840  
Db 781 EDNLVLTPEDDLCPAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840  
Qy 841 SDSSVVRGNARLPVKWMAPESLFEGYITIKSDVMSYGILLWEIFSLGWNPPGIPVDAN 900  
Db 841 SDSSVVRGNARLPVKWMAPESLFEGYITIKSDVMSYGILLWEIFSLGWNPPGIPVDAN 900  
Qy 901 FYKLIQSGFKMEOPPYATGEGYFVMSQWAFDRKRPSPNLTSLGCOLAEEACIRT 960  
Db 901 FYKLIQSGFKMEOPPYATGEGYFVMSQWAFDRKRPSPNLTSLGCOLAEEACIRT 960  
Qy 961 SIHLPKQAAPQOQGLRAQSPQVKIHRERS 992  
Db 961 SIHLPKQAAPQOQGLRAQSPQVKIHRERS 992

RESULT 5  
US-07-906-397A-2  
; Sequence 2, Application US/07906397A  
; Patent No. 5621090  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMCLONE SYSTEMS INCORPORATED  
; STREET: 180 VARICK STREET  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/906,397A  
; FILING DATE: 19920626  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/813,593  
; FILING DATE: 24-DEC-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/793,065  
; FILING DATE: 15-NOV-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/728,913  
; FILING DATE: 28-JUN-1991

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/679,666  
; FILING DATE: 02-APR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Feit, Irving N.  
; REGISTRATION NUMBER: 28,601  
; REFERENCE/DOCKET NUMBER: LEM-3-PPPPPP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-645-1405  
; TELEFAX: 212-645-2054  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 992 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-07-906-397A-2

Query Match 100.0%; Score 5264; DB 1; Length 992;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRALAQSDRRLLLVLSVMILETVTNQDLPVVKCVLISHENNGSSAGKPPSSYRVRGS 60  
Db 1 MRALAQSDRRLLLVLSVMILETVTNQDLPVVKCVLISHENNGSSAGKPPSSYRVRGS 60  
Qy 61 PEDLQCTPRROSEGTVYEAATVEVAESGSIITLQVQLATPGDLSCLVWFKHSSILGCQPHFD 120  
Db 61 PEDLQCTPRROSEGTVYEAATVEVAESGSIITLQVQLATPGDLSCLVWFKHSSILGCQPHFD 120  
Qy 121 LQNRGIVSMAILNVTTQAGEYLLHIQSERANYTVLFTVNVDRDTQLYLRRPFRKMEQ 180  
Db 121 LQNRGIVSMAILNVTTQAGEYLLHIQSERANYTVLFTVNVDRDTQLYLRRPFRKMEQ 180  
Qy 181 DALLCISSEGVPEPTVEWVLCSSHRESCKEEGPAVVRKEKVLHELFGTDIRCCARNALGR 240  
Db 181 DALLCISSEGVPEPTVEWVLCSSHRESCKEEGPAVVRKEKVLHELFGTDIRCCARNALGR 240  
Qy 241 ECTKLTIDLNQAPQSTLPQLFKYGEPLWTRCKAIHVNHGFLTWLEDEKALEGSYFE 300  
Db 241 ECTKLTIDLNQAPQSTLPQLFKYGEPLWTRCKAIHVNHGFLTWLEDEKALEGSYFE 300  
Qy 301 MSTYSTNRMTIRILLAFVSSVGRNDTGYTCCSSKHPQSALVTILEKGFINATSSQEEY 360  
Db 301 MSTYSTNRMTIRILLAFVSSVGRNDTGYTCCSSKHPQSALVTILEKGFINATSSQEEY 360  
Qy 361 EIDPYEKFCFSVRPKAYPRIRCTWTIPSOASPPCEQORGLDGYISKFCDHKKKPGHYIY 420  
Db 361 EIDPYEKFCFSVRPKAYPRIRCTWTIPSOASPPCEQORGLDGYISKFCDHKKKPGHYIY 420  
Qy 421 AENDDAQFTKMTFLNIRKKPQVLNANASQASCSGYPPLPSWTWKCKSDKSPNCTEIP 480  
Db 421 AENDDAQFTKMTFLNIRKKPQVLNANASQASCSGYPPLPSWTWKCKSDKSPNCTEIP 480  
Qy 481 EGVNKKANRKYFGQWSSSTLNSEAGKLLVKKCCAYNSMGTSCTETFLNSGPPPTQ 540  
Db 481 EGVNKKANRKYFGQWSSSTLNSEAGKLLVKKCCAYNSMGTSCTETFLNSGPPPTQ 540  
Qy 541 DNISFYATIGLCPPIVVLIVLICHYKQKQFYESQLOMIQVTGPDNDNEYFYVDFRDY 600  
Db 541 DNISFYATIGLCPPIVVLIVLICHYKQKQFYESQLOMIQVTGPDNDNEYFYVDFRDY 600  
Qy 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISTGVSIOVAVKMLKEKADSCKEALM 660  
Db 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISTGVSIOVAVKMLKEKADSCKEALM 660  
Qy 661 SELKQWTHLGHNDINVLGACTLSPVYLIIEYCCYGDLLNLSKREKFRHTWTEIPK 720  
Db 661 SELKQWTHLGHNDINVLGACTLSPVYLIIEYCCYGDLLNLSKREKFRHTWTEIPK 720  
Qy 721 EHNFSYPTFOAHNSNMPGSRVQLHPPDLQSGFNGNSIHSDEIEYENOKRLAEERE 780  
Db 721 EHNFSYPTFOAHNSNMPGSRVQLHPPDLQSGFNGNSIHSDEIEYENOKRLAEERE 780



QY 781 EDNLVLTEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKVICKDFGLARDIL 840  
DB 781 EDNLVLTEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKVICKDFGLARDIL 840  
QY 841 SDSSVVRGNARLPVKWMAPESLFGIYTIKSDVMSYGLLLWEIFSLGNVPYPGIPVDAN 900  
DB 841 SDSSVVRGNARLPVKWMAPESLFGIYTIKSDVMSYGLLLWEIFSLGNVPYPGIPVDAN 900  
QY 901 FYKLIQSGFKMEQPFVATGFIYVQSCWAFDSRKRPPNLTSLFGCOLAABEACIRT 960  
DB 901 FYKLIQSGFKMEQPFVATGFIYVQSCWAFDSRKRPPNLTSLFGCOLAABEACIRT 960  
QY 961 SIHLFKQAAPQORGGLRAQSPQVKIHRERS 992  
DB 961 SIHLFKQAAPQORGGLRAQSPQVKIHRERS 992

RESULT 6

US-08-601-891-2  
Sequence 2, Application US/08601891  
Patent No. 5747651  
GENERAL INFORMATION:  
APPLICANT: Lemischka, Ihor R.  
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESS: Imclone Systems Incorporated  
STREET: 180 Varick Street  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10014  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/601,891  
FILING DATE: 15-FEB-1996  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/977,451  
FILING DATE: 19-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/906,397  
FILING DATE: 26-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US PCT/US92/05401  
FILING DATE: 26-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: TW 81102961  
FILING DATE: 15-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US PCT/US92/02750  
FILING DATE: 02-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/813,593  
FILING DATE: 24-DEC-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/793,065  
FILING DATE: 15-NOV-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/728,913  
FILING DATE: 28-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Feit, Irving N.  
REGISTRATION NUMBER: 28,601

REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 992 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-601-891-2

Query Match 100.0%; Score 5264; DB 1; Length 992;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALAQRSDRRLLLVLSVMILEVTNQDLPVKICVLISHENNSSACKPSYSYRMVRS 60  
DB 1 MRALAQRSDRRLLLVLSVMILEVTNQDLPVKICVLISHENNSSACKPSYSYRMVRS 60  
QY 61 PEDLQCTPRQSEGTVEAATVEVAESGSITLQVQLATPGDLSCLWVFKHSLGCPHF 120  
DB 61 PEDLQCTPRQSEGTVEAATVEVAESGSITLQVQLATPGDLSCLWVFKHSLGCPHF 120  
QY 121 LQNRGIVSMALNVTTQAGEYLLHIQSERANYTVLFTVNVDRDTQLYVLRPFYRQWENQ 180  
DB 121 LQNRGIVSMALNVTTQAGEYLLHIQSERANYTVLFTVNVDRDTQLYVLRPFYRQWENQ 180  
QY 181 DALLCISGVPEPTVEWVLCSSHRESCKEGBPAVVRKEKVLHELPGTDIRCCARNALGR 240  
DB 181 DALLCISGVPEPTVEWVLCSSHRESCKEGBPAVVRKEKVLHELPGTDIRCCARNALGR 240  
QY 241 ECTKLTITDLNQAPOSTLPQLFKVGEPLWIRCKAIHVNHGFGLTWELEDKALEGSYFE 300  
DB 241 ECTKLTITDLNQAPOSTLPQLFKVGEPLWIRCKAIHVNHGFGLTWELEDKALEGSYFE 300  
QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCSHSPQSALVTLKGFINATSQEY 360  
DB 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCSHSPQSALVTLKGFINATSQEY 360  
QY 361 EIDPYEKFCFSVRKAYPRICTWIFSOASFFCEORGLDGYISIKFCDHKKPGYIFY 420  
DB 361 EIDPYEKFCFSVRKAYPRICTWIFSOASFFCEORGLDGYISIKFCDHKKPGYIFY 420  
QY 421 AENDDAQFTKMTNIRKPKQVLANASQAQSSDGYPLPSWTWKCKSDKSPNCTEIP 480  
DB 421 AENDDAQFTKMTNIRKPKQVLANASQAQSSDGYPLPSWTWKCKSDKSPNCTEIP 480  
QY 481 EGVWNNKANRKFVGQVSSSTLNMSEAGKGLLVKCCAYNSMGTSCTIFLNSPGPPFIQ 540  
DB 481 EGVWNNKANRKFVGQVSSSTLNMSEAGKGLLVKCCAYNSMGTSCTIFLNSPGPPFIQ 540  
QY 541 DNISFYATIGLCLPFIIVLILVLI CHYKQKQFYESQLOMIQVTPGLDNEYFYVDFRDY 600  
DB 541 DNISFYATIGLCLPFIIVLILVLI CHYKQKQFYESQLOMIQVTPGLDNEYFYVDFRDY 600  
QY 601 DLKWEFPRENLEFGKVLGSGAFGRVNNATAYGISKTVGSIQVAVMLKADKCEKALM 660  
DB 601 DLKWEFPRENLEFGKVLGSGAFGRVNNATAYGISKTVGSIQVAVMLKADKCEKALM 660  
QY 661 SELKMWTHLGHHDNIVNLGACTLSGPVYLI FEYCCYCGDLLNVLRSKREKFRHTWTIEFK 720  
DB 661 SELKMWTHLGHHDNIVNLGACTLSGPVYLI FEYCCYCGDLLNVLRSKREKFRHTWTIEFK 720  
QY 721 EHNFSYPTFOAHNSNMPGSRVQLHPPLDQLSGFNGNSIHSEDEIYENQRLAEEEE 780  
DB 721 EHNFSYPTFOAHNSNMPGSRVQLHPPLDQLSGFNGNSIHSEDEIYENQRLAEEEE 780  
QY 781 EDNLVLTREDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKVICKDFGLARDIL 840  
DB 781 EDNLVLTREDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKVICKDFGLARDIL 840  
QY 841 SDSSVVRGNARLPVKWMAPESLFGIYTIKSDVMSYGLLLWEIFSLGNVPYPGIPVDAN 900

Db 841 SDSSVVRGNARLPVKWAPESLFGITVIRKSDVMSYGLLWEIFSLGWNPPGIPVDAN 900  
QY 901 FYKLIQSGFKMEQPYATGEGYFVMSQWAFDSRKRPSFNLTSFLGCOLAEAEACIRT 960  
Db 901 FYKLIQSGFKMEQPYATGEGYFVMSQWAFDSRKRPSFNLTSFLGCOLAEAEACIRT 960  
QY 961 SIHLPKQAAPQOQGLRAQSPORQVKIHRERS 992  
Db 961 SIHLPKQAAPQOQGLRAQSPORQVKIHRERS 992

RESULT 7

US-09-021-324-2  
Sequence 2, Application US/09021324  
Patent No. 5912133  
GENERAL INFORMATION:  
APPLICANT: Lemischka, Ihor R.  
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ImClone Systems Incorporated  
STREET: 180 Varick Street  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10014  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/021.324  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/977.451  
FILING DATE: 1992-11-19  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/906.397  
FILING DATE: 26-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US PCT/US92/05401  
FILING DATE: 26-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: TW 81102961  
FILING DATE: 15-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US PCT/US92/02750  
FILING DATE: 02-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/813.593  
FILING DATE: 24-DEC-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/793.065  
FILING DATE: 15-NOV-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/728.913  
FILING DATE: 28-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/679.666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Felt, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEW-3-7P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:

LENGTH: 992 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-021-324-2  
Query Match 100.0%; Score 5264; DB 2; Length 992;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRALAQRSDRRLLLLVLSVMILETTVTNODLPVIAKVLISHENNGSSACKPSSYRMVGRS 60  
Db 1 MRALAQRSDRRLLLLVLSVMILETTVTNODLPVIAKVLISHENNGSSACKPSSYRMVGRS 60  
QY 61 PEDLOCTPRROSEGTVEAATVEVAESGITLOVOLATPGDLSCLVFKHSSLGCCPHFD 120  
Db 61 PEDLOCTPRROSEGTVEAATVEVAESGITLOVOLATPGDLSCLVFKHSSLGCCPHFD 120  
QY 121 LQNRGIVSMALINVTETQAGEYLLHIQSERANYTLVTVNVRDTOLYVLRPFRKWMENQ 180  
Db 121 LQNRGIVSMALINVTETQAGEYLLHIQSERANYTLVTVNVRDTOLYVLRPFRKWMENQ 180  
QY 181 DALLCISEGVPEPTVEWVLCSSHRESCKEEGPAVVRKEEKVLHELFGTDIRCCARNALGR 240  
Db 181 DALLCISEGVPEPTVEWVLCSSHRESCKEEGPAVVRKEEKVLHELFGTDIRCCARNALGR 240  
QY 241 ECTKLFTIDLNAQPOSTLPQLFKVGEPLWIRCKAIHVNHGFGLTWELEDKALEEGSYFE 300  
Db 241 ECTKLFTIDLNAQPOSTLPQLFKVGEPLWIRCKAIHVNHGFGLTWELEDKALEEGSYFE 300  
QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCCSSRHPQSALAVTILEKGFINATSSQEEY 360  
Db 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCCSSRHPQSALAVTILEKGFINATSSQEEY 360  
QY 361 EIDPYEKEFCFSVRFKAYPRICTWIFSOASPFCEORGLDGYISIKFCDHKHKKPGYIFY 420  
Db 361 EIDPYEKEFCFSVRFKAYPRICTWIFSOASPFCEORGLDGYISIKFCDHKHKKPGYIFY 420  
QY 421 AENDDAQTKMFTLNIRKPKQVLANASASQACSSDGYPLPSWTWKCKSDKSPNCTEETIP 480  
Db 421 AENDDAQTKMFTLNIRKPKQVLANASASQACSSDGYPLPSWTWKCKSDKSPNCTEETIP 480  
QY 481 EGVNKKANRKFVQWVSSSLNMSHAGKLLVKCCAYNSMGTSCTETIFLNSPGPPPTIQ 540  
Db 481 EGVNKKANRKFVQWVSSSLNMSHAGKLLVKCCAYNSMGTSCTETIFLNSPGPPPTIQ 540  
QY 541 DNISFYATIGLCPLPIVVLIVLICHYKKQFYESQLOMIQVTPGLDNEFYVVDPRDYEY 600  
Db 541 DNISFYATIGLCPLPIVVLIVLICHYKKQFYESQLOMIQVTPGLDNEFYVVDPRDYEY 600  
QY 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISKTGVSIVQAVKMLKEKADSCKEALM 660  
Db 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISKTGVSIVQAVKMLKEKADSCKEALM 660  
QY 661 SELKQWTHLGHNDINVLGACTLSGPVYLIFEYCCYGDLLNLYLRSKREKFKHRTWTEIFK 720  
Db 661 SELKQWTHLGHNDINVLGACTLSGPVYLIFEYCCYGDLLNLYLRSKREKFKHRTWTEIFK 720  
QY 721 EBNFSSYPTFOAHNSNMPGSRREVQLHPDLQSGFNCSHSEDEIEYENOKRLAEEREE 780  
Db 721 EBNFSSYPTFOAHNSNMPGSRREVQLHPDLQSGFNCSHSEDEIEYENOKRLAEEREE 780  
QY 781 EDLNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGVKVKICDFGLARDIL 840  
Db 781 EDLNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGVKVKICDFGLARDIL 840  
QY 841 SDSSVVRGNARLPVKWAPESLFGITVIRKSDVMSYGLLWEIFSLGWNPPGIPVDAN 900  
Db 841 SDSSVVRGNARLPVKWAPESLFGITVIRKSDVMSYGLLWEIFSLGWNPPGIPVDAN 900  
QY 901 FYKLIQSGFKMEQPYATGEGYFVMSQWAFDSRKRPSFNLTSFLGCOLAEAEACIRT 960  
Db 901 FYKLIQSGFKMEQPYATGEGYFVMSQWAFDSRKRPSFNLTSFLGCOLAEAEACIRT 960

QY 961 SIHLPKQAAPQORGLRAQSPQORQVHIHRS 992  
Db 961 SIHLPKQAAPQORGLRAQSPQORQVHIHRS 992

RESULT 8

PCT-US92-02750-2

; Sequence 2, Application PC/TUS9202750

; GENERAL INFORMATION:

; APPLICANT: LEMISCHKA, ITHOR R.

; TITLE OF INVENTION: Totipotent Hematopoietic Stem Cell

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: IMCLONE SYSTEMS INCORPORATED

; STREET: 180 VARICK STREET

; CITY: NEW YORK

; STATE: NEW YORK

; COUNTRY: US

; ZIP: 10014

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US92/02750

; FILING DATE: 19920402

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: FEIT, IRVING N.

; REGISTRATION NUMBER: 28,601

; REFERENCE/DOCKET NUMBER: LEM-3-PPPT

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-645-1405

; TELEFAX: 212-645-2054

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 992 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

; MOLECULE TYPE: protein

PCT-US92-02750-2

Query Match 100.0%; Score 5264; DB 5; Length 992;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALQSRDRRLLLVLSVMILETVTNQDLPVVKVLSHENNGSSAGKPSYRMVRS 60

Db 1 MRALQSRDRRLLLVLSVMILETVTNQDLPVVKVLSHENNGSSAGKPSYRMVRS 60

QY 61 PEDLQTPRQSEGTVEAATVEAABSGSITLQVQLATPGDLSCVLWFKHSLGCOHPD 120

Db 61 PEDLQTPRQSEGTVEAATVEAABSGSITLQVQLATPGDLSCVLWFKHSLGCOHPD 120

QY 121 LQNRGIVSNAILNVTQAGEVLLHIQSERANYTLFTVNVNRTDQLYVLRPFKMNQ 180

Db 121 LQNRGIVSNAILNVTQAGEVLLHIQSERANYTLFTVNVNRTDQLYVLRPFKMNQ 180

QY 181 DALLCISEGVPEPTVEWVLCSSHRESCKEKGPAVRKEKVLHELFGTDIRCARNALGR 240

Db 181 DALLCISEGVPEPTVEWVLCSSHRESCKEKGPAVRKEKVLHELFGTDIRCARNALGR 240

QY 241 ECTKLTIDLNQAPSTLPQLFLKVGEPWIRCKAIHVNHGFLTWLEDKALEGSYFE 300

Db 241 ECTKLTIDLNQAPSTLPQLFLKVGEPWIRCKAIHVNHGFLTWLEDKALEGSYFE 300

QY 301 MSTYGTNRTMIRILLAFVSSVGRNDTGYTCCSSKHPQSALVTILEKGFINATSSQEEY 360

Db 301 MSTYGTNRTMIRILLAFVSSVGRNDTGYTCCSSKHPQSALVTILEKGFINATSSQEEY 360

QY 361 EIDPYEKFCSVRFKAYPRIRCTWIFSQAQSPFCEORGLGLEDGYSISKFDHKNKPGYIFY 420  
Db 361 EIDPYEKFCSVRFKAYPRIRCTWIFSQAQSPFCEORGLGLEDGYSISKFDHKNKPGYIFY 420  
QY 421 AENDDAQFTKMTLNIRKKPQVLANASASQASCSGDPPLPSWTWKCKSDKSPNCTEEIP 480  
Db 421 AENDDAQFTKMTLNIRKKPQVLANASASQASCSGDPPLPSWTWKCKSDKSPNCTEEIP 480  
QY 481 EGVNKKANRVFGQWVSSTLNMSEAGKGLLVKCCAYNSMGTSCETIFLNSPGPPFFIQ 540  
Db 481 EGVNKKANRVFGQWVSSTLNMSEAGKGLLVKCCAYNSMGTSCETIFLNSPGPPFFIQ 540  
QY 541 DNISFYATIGLCLPIVVLVILCHYKQKQFYESQLOMIQVTGPLDNEFYVDFRDY 600  
Db 541 DNISFYATIGLCLPIVVLVILCHYKQKQFYESQLOMIQVTGPLDNEFYVDFRDY 600  
QY 601 DLKWEFFRENLEFGKVLGSGAFGRVMNATAYGISTGTGYSIOAVKMLKEKADSCKEALM 660  
Db 601 DLKWEFFRENLEFGKVLGSGAFGRVMNATAYGISTGTGYSIOAVKMLKEKADSCKEALM 660  
QY 661 SELKXWTHLGHNDINVLIGACTLSGPVYLIPEYCCYGDLLNLYLSKREKFRHTWTEIFK 720  
Db 661 SELKXWTHLGHNDINVLIGACTLSGPVYLIPEYCCYGDLLNLYLSKREKFRHTWTEIFK 720  
QY 721 EHNFSYPTFOAHSNSMPGSRVQLHPPLDQLSGFNNGSIHSEDEIEYENQRLAEBEE 780  
Db 721 EHNFSYPTFOAHSNSMPGSRVQLHPPLDQLSGFNNGSIHSEDEIEYENQRLAEBEE 780  
QY 781 EDLNVLTFTEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840  
Db 781 EDLNVLTFTEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840  
QY 841 SDSSVVVRGNARLPVKWMAPELSLPGIYTIKSDVMSYGILLWEIFSLGYNPYGIPVDAN 900  
Db 841 SDSSVVVRGNARLPVKWMAPELSLPGIYTIKSDVMSYGILLWEIFSLGYNPYGIPVDAN 900  
QY 901 FYKLIQSGFKMEQPFYATGTYFVMSQWAFDRSRKPSFPNLTSLFGCOLAEBEACIRT 960  
Db 901 FYKLIQSGFKMEQPFYATGTYFVMSQWAFDRSRKPSFPNLTSLFGCOLAEBEACIRT 960  
QY 961 SIHLPKQAAPQORGLRAQSPQORQVHIHRS 992  
Db 961 SIHLPKQAAPQORGLRAQSPQORQVHIHRS 992

RESULT 9

PCT-US92-05401-2

; Sequence 2, Application PC/TUS9205401

; GENERAL INFORMATION:

; APPLICANT: Lemischka, Ithor R.

; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL

; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: IMCLONE SYSTEMS INCORPORATED

; STREET: 180 VARICK STREET

; CITY: NEW YORK

; STATE: NEW YORK

; COUNTRY: U.S.A.

; ZIP: 10014

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US92/05401

; FILING DATE: 19920626

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Feit, Irving N.

; REGISTRATION NUMBER: 28,601

; REFERENCE/DOCKET NUMBER: LEM-3-PPPPPT

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 992 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US92-05401-2

Query Match 100.0%; Score 5264; DB 5; Length 992;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALQSRDRRLLLVLSVMIETVNDLPVVKVLSHENNGSSAGKPSYRMVRS 60  
DB 1 MRALQSRDRRLLLVLSVMIETVNDLPVVKVLSHENNGSSAGKPSYRMVRS 60

QY 61 PEDLOCTPRQSEGTVEAATVEAASGSITLQVQLATPGDLSCLVFKHSSLGCCQPHD 120  
DB 61 PEDLOCTPRQSEGTVEAATVEAASGSITLQVQLATPGDLSCLVFKHSSLGCCQPHD 120

QY 121 LQNRGIVSMALNVTTQAGEYLLHIQSERANTVLTFTVNVRTQLYVLRPFYRKWENQ 180  
DB 121 LQNRGIVSMALNVTTQAGEYLLHIQSERANTVLTFTVNVRTQLYVLRPFYRKWENQ 180

QY 181 DALLCISEGVPEPTVEWVLCSSHRESCKEKGPAVVRKEEKLHELFGTDIRCCARNALGR 240  
DB 181 DALLCISEGVPEPTVEWVLCSSHRESCKEKGPAVVRKEEKLHELFGTDIRCCARNALGR 240

QY 241 ECTKLTIDLNOAQPOSTLPQLFKVGEPLWIRCKAIHVNHGFLTWELDKALEEGSYFE 300  
DB 241 ECTKLTIDLNOAQPOSTLPQLFKVGEPLWIRCKAIHVNHGFLTWELDKALEEGSYFE 300

QY 301 MSTYSTNRTMIRLLAFVSSVGRNDGYTCSSSKHPSQSALVTILEKGFINATSSQEEY 360  
DB 301 MSTYSTNRTMIRLLAFVSSVGRNDGYTCSSSKHPSQSALVTILEKGFINATSSQEEY 360

QY 361 EIDPYEKFCFSVFKAYPIRCTWIFPSQAFPCOEQRLEDGYISKFDHKNKPGYIFY 420  
DB 361 EIDPYEKFCFSVFKAYPIRCTWIFPSQAFPCOEQRLEDGYISKFDHKNKPGYIFY 420

QY 421 AENDDAQFTKMTLNIRKPKQVLANASASQASCSGYPPLPSMTWKCKDKSPNCTEIP 480  
DB 421 AENDDAQFTKMTLNIRKPKQVLANASASQASCSGYPPLPSMTWKCKDKSPNCTEIP 480

QY 481 EGVNKKANRKGQWVSSSTLNMSEAGKLLVKKCCAYNSMGTSCTIFLNSPGPPPTIQ 540  
DB 481 EGVNKKANRKGQWVSSSTLNMSEAGKLLVKKCCAYNSMGTSCTIFLNSPGPPPTIQ 540

QY 541 DNISFYATIGLCLPFTIWLIVLICHVKYKQFVYESQLOMIQVTPGLDNEFYVDPEDYEV 600  
DB 541 DNISFYATIGLCLPFTIWLIVLICHVKYKQFVYESQLOMIQVTPGLDNEFYVDPEDYEV 600

QY 601 DLKWEFFRENLEBFGKVLGSGAFGRVNMATAYGISTGVSIQAVAKMLKEKADSCEKEALM 660  
DB 601 DLKWEFFRENLEBFGKVLGSGAFGRVNMATAYGISTGVSIQAVAKMLKEKADSCEKEALM 660

QY 661 SELKMWTHLGHNDIVNLGACTLSGPVYLIIEYCCYGDLLNLYLSKRSKRFHRTWTEIPK 720  
DB 661 SELKMWTHLGHNDIVNLGACTLSGPVYLIIEYCCYGDLLNLYLSKRSKRFHRTWTEIPK 720

QY 721 ENHFSYPTFOAHSNMPGSEVQLHPPLDOLSGFNGNSIHSEDEIEVENOKRLAEERE 780  
DB 721 ENHFSYPTFOAHSNMPGSEVQLHPPLDOLSGFNGNSIHSEDEIEVENOKRLAEERE 780

QY 781 EDNLVLTPEDLFCFAYQVAKMEFLEFKSCVHRDLAARNVLTGHGVKVICDFGLARDIL 840  
DB 781 EDNLVLTPEDLFCFAYQVAKMEFLEFKSCVHRDLAARNVLTGHGVKVICDFGLARDIL 840

QY 841 SDSSYVVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIFSLGVNYPYGPVVDAN 900

DB 841 SDSSYVVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIFSLGVNYPYGPVVDAN 900  
QY 901 FYKLIQSGFKMEQPFYATGFIYFVMSQWAFDSRKPSPNLTSPGLQCLAEAEACIRT 960  
DB 901 FYKLIQSGFKMEQPFYATGFIYFVMSQWAFDSRKPSPNLTSPGLQCLAEAEACIRT 960  
QY 961 SIHLPKQAAPQORGGLRAQSPQOVKIHRS 992  
DB 961 SIHLPKQAAPQORGGLRAQSPQOVKIHRS 992

RESULT 10

PCT-US92-09893-2  
Sequence 2, Application PC/TUS9209893  
GENERAL INFORMATION:  
APPLICANT: Lemischka, Ihor R.  
TITLE OF INVENTION: TOTIPOENT HEMATOPOIETIC STEM CELL  
TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ImClone Systems Incorporated  
STREET: 180 Varick Street  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10014  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/09893  
FILING DATE: 19921116  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Feit, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-7PT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 992 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US92-09893-2

Query Match 100.0%; Score 5264; DB 5; Length 992;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALQSRDRRLLLVLSVMIETVNDLPVVKVLSHENNGSSAGKPSYRMVRS 60  
DB 1 MRALQSRDRRLLLVLSVMIETVNDLPVVKVLSHENNGSSAGKPSYRMVRS 60

QY 61 PEDLOCTPRQSEGTVEAATVEAASGSITLQVQLATPGDLSCLVFKHSSLGCCQPHD 120  
DB 61 PEDLOCTPRQSEGTVEAATVEAASGSITLQVQLATPGDLSCLVFKHSSLGCCQPHD 120

QY 121 LQNRGIVSMALNVTTQAGEYLLHIQSERANTVLTFTVNVRTQLYVLRPFYRKWENQ 180  
DB 121 LQNRGIVSMALNVTTQAGEYLLHIQSERANTVLTFTVNVRTQLYVLRPFYRKWENQ 180

QY 181 DALLCISEGVPEPTVEWVLCSSHRESCKEKGPAVVRKEEKLHELFGTDIRCCARNALGR 240  
DB 181 DALLCISEGVPEPTVEWVLCSSHRESCKEKGPAVVRKEEKLHELFGTDIRCCARNALGR 240

QY 241 ECTKLTIDLNOAQPOSTLPQLFKVGEPLWIRCKAIHVNHGFLTWELDKALEEGSYFE 300  
DB 241 ECTKLTIDLNOAQPOSTLPQLFKVGEPLWIRCKAIHVNHGFLTWELDKALEEGSYFE 300

```

301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTTCSSSKHPSQSALVTILEKGFINATSSQEEY 360
301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTTCSSSKHPSQSALVTILEKGFINATSSQEEY 360
361 EIDPYEKFCSVRFKAYPRIRCTWIPFQASFPCEQRLGLEDGYSISKFCDHKNKPGYIFY 420
361 EIDPYEKFCSVRFKAYPRIRCTWIPFQASFPCEQRLGLEDGYSISKFCDHKNKPGYIFY 420
421 AENDDAQFTKMTLIRKPKQVLANASASQASCSDDGYPLPSWTWKCDKSPNCTEEIP 480
421 AENDDAQFTKMTLIRKPKQVLANASASQASCSDDGYPLPSWTWKCDKSPNCTEEIP 480
481 EGVNKKANRKYFGQVSSSTLNMSEAGKGLLVKCCAYNSMGTSCTIFLNSPGPPFFIQ 540
481 EGVNKKANRKYFGQVSSSTLNMSEAGKGLLVKCCAYNSMGTSCTIFLNSPGPPFFIQ 540
541 DNISFYATTIGLCLPFIWLVILCHIKYKQFYESQLOMIQVTPGLDNEYFYVDFRDY 600
541 DNISFYATTIGLCLPFIWLVILCHIKYKQFYESQLOMIQVTPGLDNEYFYVDFRDY 600
601 DLKWEPPRENLEFGKVLGSGAFGRVNNATAYGISKTVSTIOVAVKMLKEKADSCEKALM 660
601 DLKWEPPRENLEFGKVLGSGAFGRVNNATAYGISKTVSTIOVAVKMLKEKADSCEKALM 660
661 SELKMWTHLGHNDINVLGACTLSGPVYLIFEYCCYGDLLNLYLRSKREKFRHTWTEIFK 720
661 SELKMWTHLGHNDINVLGACTLSGPVYLIFEYCCYGDLLNLYLRSKREKFRHTWTEIFK 720
721 EHNFSYPTFOAHNSMPCGSEVQLHPPDLQSLGPNNGSIHSEDEIEYENQKLAEEBE 780
721 EHNFSYPTFOAHNSMPCGSEVQLHPPDLQSLGPNNGSIHSEDEIEYENQKLAEEBE 780
781 EDNLVLTPEDLICFAQVAKGMEFLFKSCVHRDLAARNVLTGHGVKVICDFGLARDIL 840
781 EDNLVLTPEDLICFAQVAKGMEFLFKSCVHRDLAARNVLTGHGVKVICDFGLARDIL 840
841 SDSSYVVRGNARLPVKWMAPELFEIGIYIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900
841 SDSSYVVRGNARLPVKWMAPELFEIGIYIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900
901 FYKLIQSGFKMOPPFVATGIVFMQSCWAFSRKPSFPNLTSTFLGCOLAEEACIRT 960
901 FYKLIQSGFKMOPPFVATGIVFMQSCWAFSRKPSFPNLTSTFLGCOLAEEACIRT 960
961 SIHLPKQAAPQORGGIARAQSPORQVXIHRERS 992
961 SIHLPKQAAPQORGGIARAQSPORQVXIHRERS 992

RESULT 11
US-08-222-299-2
; Sequence 2, Application US/08222299
; Patent No. 5635388
; GENERAL INFORMATION:
; APPLICANT: Bennett, Brian D.
; APPLICANT: Broz, Susan D.
; APPLICANT: Matthews, William
; APPLICANT: Zeigler, Francis C.
; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES THEREIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)

```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,299
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Hasak, Janet E.
REGISTRATION NUMBER: 28,616
REFERENCE/DOCKET NUMBER: 879
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1896
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1000 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-222-299-2

Query Match 96.9%; Score 5102; DB 1; Length 1000;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 968; Conservative 6; Mismatches 17; Indels 10; Gaps 2;

QY 1 MRALQSRDRRLLLVLSVMIETVTNODLPVKVCLISHENNGSSAGKPPSSYRMVRGS 60
DB 1 MRALQSRDRRLLLVLSVMIETVTNODLPVKVCLISHENNGSSAGKPPSSYRMVRGS 60
QY 61 PEDLOCTPRQSEGVYEAATVEAESGSITLQVQLATPGDLSCLVFKHSSLGCPHF 120
DB 61 PEDLOCTPRQSEGVYEAATVEAESGSITLQVQLATPGDLSCLVFKHSSLGCPHF 120
QY 121 LQNRGIVSAIINVTETQAGEYLLHIQSERANYTLFTVNRDTQIYVLRPFRKME 180
DB 121 LQNRGIVSAIINVTETQAGEYLLHIQSERANYTLFTVNRDTQIYVLRPFRKME 180
QY 181 DALLCISEGVPTVEWVLCSSHRSCKEEGPAVVRKEKVLHELPGTDIRCCARNALGR 240
DB 181 DALLCISEGVPTVEWVLCSSHRSCKEEGPAVVRKEKVLHELPGTDIRCCARNALGR 240
QY 241 ECTKLTIDLNQAPOSTLPQLFKVGEPLWIRCKAIHVNHGFLTWELDKALEEGSYFE 300
DB 241 ECTKLTIDLNQAPOSTLPQLFKVGEPLWIRCKAIHVNHGFLTWELDKALEEGSYFE 300
QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTTCSSSKHPSQSALVTILEKGFINATSSQEEY 360
DB 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTTCSSSKHPSQSALVTILEKGFINATSSQEEY 360
QY 361 EIDPYEKFCSVRFKAYPRIRCTWIPFQASFPCEQRLGLEDGYSISKFCDHKNKPGYIFY 420
DB 361 EIDPYEKFCSVRFKAYPRIRCTWIPFQASFPCEQRLGLEDGYSISKFCDHKNKPGYIFY 420
QY 421 AENDDAQFTKMTLIRKPKQVLANASASQASCSDDGYPLPSWTWKCDKSPNCTEEIP 480
DB 421 AENDDAQFTKMTLIRKPKQVLANASASQASCSDDGYPLPSWTWKCDKSPNCTEEIP 480
QY 481 EGVNKKANRKYFGQVSSSTLNMSEAGKGLLVKCCAYNSMGTSCTIFLNSPGPPFFIQ 540
DB 481 EGVNKKANRKYFGQVSSSTLNMSEAGKGLLVKCCAYNSMGTSCTIFLNSPGPPFFIQ 540
QY 541 DNISFYATTIGLCLPFIWLVILCHIKYKQFYESQLOMIQVTPGLDNEYFYVDFRDY 600
DB 541 DNISFYATTIGLCLPFIWLVILCHIKYKQFYESQLOMIQVTPGLDNEYFYVDFRDY 600
QY 601 DLKWEPPRENLEFGKVLGSGAFGRVNNATAYGISKTVSTIOVAVKMLKEKADSCEKALM 660
DB 601 DLKWEPPRENLEFGKVLGSGAFGRVNNATAYGISKTVSTIOVAVKMLKEKADSCEKALM 660
QY 661 SELKMWTHLGHNDINVLGACTLSGPVYLIFEYCCYGDLLNLYLRSKREKFRHTWTEIFK 720
DB 661 SELKMWTHLGHNDINVLGACTLSGPVYLIFEYCCYGDLLNLYLRSKREKFRHTWTEIFK 720

```

Qy 721 EHNFSYPTFOAHNSMPSGSEVQLHPPLDQLSGFNGNSIHSEDEIEYENQKRLAEBEE 780  
 Db 721 EHNFSYPTFOAHNSMPSGSEVQLHPPLDQLSGFNGNSIHSEDEIEYENQKRLAEBEE 780  
 Qy 781 EDNLVTFEDLLCFAYQVAKMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840  
 Db 781 EDNLVTFEDLLCFAYQVAKMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840  
 Qy 841 SDSSVVRGNARLPVKWMAPELFEIYIKSDVMSYGILLWEIFSLGWNVPYGPVVDAN 900  
 Db 841 SDSSVVRGNARLPVKWMAPELFEIYIKSDVMSYGILLWEIFSLGWNVPYGPVVDAN 900  
 Qy 901 FYKLIQSGFKMEQPFYATEGIYFVMSQWAFDSRKRPSFPNLTSLGQLAEAEACIRT 960  
 Db 901 FYKLIQSGFKMEQPFYATEGIYFVMSQWAFDSRKRPSFPNLTSLGQLAEAEACIRT 960  
 Qy 961 -----S1HLPKQAAPOQRGGLRAQSPORQVKIHRERS 992  
 Db 961 MGNVPEHPSIQNRRLPSREAGS-EPPSPQAQVKIHRERS 1000

RESULT 12

US-08-434-878-2

; Sequence 2, Application US/08434878

; Patent No. 5997865

; GENERAL INFORMATION:

; APPLICANT: Bennett, Brian D.

; APPLICANT: Broz, Susan D.

; APPLICANT: Matthews, William

; APPLICANT: Zeigler, Francis C.

; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES THEREOF

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.

; STREET: 460 Point San Bruno Blvd

; CITY: South San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94080

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patin (Genentech)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/434,878

; FILING DATE:

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Hasak, Janet E.

; REGISTRATION NUMBER: 28,616

; REFERENCE/DOCKET NUMBER: 879

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415/225-1896

; TELEFAX: 415/952-9881

; TELEX: 910/371-7168

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1000 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

US-08-434-878-2

Query Match 96.9%; Score 5102; DB 2; Length 1000;

Best Local Similarity 96.7%; Pred. No. 0;

Matches 968; Conservative 6; Mismatches 17; Indels 10; Gaps 2;

Qy 1 MRALAQRSDRLLLVLSVMILETTVNQDLPIVKCVLISHENNGSSAGKPSRYMVRGS 60  
 |||||

Db 1 MRALAQRSDRLLLVLSVMILETTVNQDLPIVKCVLISHENNGSSAGKPSRYMVRGS 60  
 Qy 61 PEDLOCTPRRQSEGTVYEAATVEVAESGSIITLQVQLATPGDLSCLVFKHSSIGCOPHFD 120  
 Db 61 PEDLOCTPRRQSEGTVYEAATVEVAESGSIITLQVQLATPGDLSCLVFKHSSIGCOPHFD 120  
 Qy 121 LQNGIIVSMALNTVETQAGEYLLHIQSERANYTVLFTVNVVDTQIYVLRVRRYFRKQENQ 180  
 Db 121 LQNGIIVSMALNTVETQAGEYLLHIQSERANYTVLFTVNVVDTQIYVLRVRRYFRKQENQ 180  
 Qy 181 DALLCISEGVPEPTVWVLCSSHRESCKEKGPAVVRKEKVLHFGTDIRCARNALGR 240  
 Db 181 DALLCISEGVPEPTVWVLCSSHRESCKEKGPAVVRKEKVLHFGTDIRCARNALGR 240  
 Qy 241 ECTKLFTIDLNQAPQSTLPQLFLKVGPELWIRCKAIHVNHGFGLTWELEDKALEGSGYPE 300  
 Db 241 ESTLFTIDLNQAPQSTLPQLFLKVGPELWIRCKAIHVNHGFGLTWELEDKALEGSGYPE 300  
 Qy 301 MSTYSTNRMTIRIILAFVSSVGRNDTGYTTCSSSKHPSQSALVTILEKGFINATSSQBEY 360  
 Db 301 MSTYSTNRMTIRIILAFVSSVGRNDTGYTTCSSSKHPSQSALVTILEKGFINATSSQBEY 360  
 Qy 361 EIDPYEKFCSVRPKAYPRIRCTWIFSOASPCQORGLDGYISIKPCDHKNKPGEYIFY 420  
 Db 361 EIDPYEKFCSVRPKAYPRIRCTWIFSOASPCQORGLDGYISIKPCDHKNKPGEYIFY 420  
 Qy 421 AENDDAQFTKMTLNIRKPKQVLANASASQASCSGYPPLPSWTWKKCSKSPNCTEIP 480  
 Db 421 AENDDAQFTKMTLNIRKPKQVLANASASQASCSGYPPLPSWTWKKCSKSPNCTEIP 480  
 Qy 481 EGVNKKANRKFQGVWSSSTLNMSEAGKLLVCCAYNSMGTSCTEIFLNSPFPPIQ 540  
 Db 481 EGVNKKANRKFQGVWSSSTLNMSEAGKLLVCCAYNSMGTSCTEIFLNSPFPPIQ 540  
 Qy 541 DNISFYATIGLCPLFTVILVILCHYKQKQFVESQLOMTQVPLONEYPYVDFRDVEY 600  
 Db 541 DNISFYATIGLCPLFTVILVILCHYKQKQFVESQLOMTQVPLONEYPYVDFRDVEY 600  
 Qy 601 DLKWEFPRENLEFGKVLGSGAPGRVMNATAYGISTGVSIOQAVKMLKEKADSCKEALM 660  
 Db 601 DLKWEFPRENLEFGKVLGSGAPGRVMNATAYGISTGVSIOQAVKMLKEKADSCKEALM 660  
 Qy 661 SELKMTLGHNDINVLGACTLSGPFVYLIFBYCCYGDLLNLYRSKREKPHRTWTIFK 720  
 Db 661 SELKMTLGHNDINVLGACTLSGPFVYLIFBYCCYGDLLNLYRSKREKPHRTWTIFK 720  
 Qy 721 EHNFSYPTFOAHNSMPSGSEVQLHPPLDQLSGFNGNSIHSEDEIEYENQKRLAEBEE 780  
 Db 721 EHNFSYPTFOAHNSMPSGSEVQLHPPLDQLSGFNGNSIHSEDEIEYENQKRLAEBEE 780  
 Qy 781 EDNLVTFEDLLCFAYQVAKMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840  
 Db 781 EDNLVTFEDLLCFAYQVAKMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840  
 Qy 841 SDSSVVRGNARLPVKWMAPELFEIYIKSDVMSYGILLWEIFSLGWNVPYGPVVDAN 900  
 Db 841 SDSSVVRGNARLPVKWMAPELFEIYIKSDVMSYGILLWEIFSLGWNVPYGPVVDAN 900  
 Qy 901 FYKLIQSGFKMEQPFYATEGIYFVMSQWAFDSRKRPSFPNLTSLGQLAEAEACIRT 960  
 Db 901 FYKLIQSGFKMEQPFYATEGIYFVMSQWAFDSRKRPSFPNLTSLGQLAEAEACIRT 960  
 Qy 961 -----S1HLPKQAAPOQRGGLRAQSPORQVKIHRERS 992  
 Db 961 MGNVPEHPSIQNRRLPSREAGS-EPPSPQAQVKIHRERS 1000

RESULT 13

PCT-US95-03718-2

; Sequence 2, Application PC/TUS9503718

; GENERAL INFORMATION:

; APPLICANT: GENENTECH, INC.

; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES THEREOF

NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/03718  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Wendy M. Lee  
REGISTRATION NUMBER: 00,000  
REFERENCE/DOCKET NUMBER: 879PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1000 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
PCT-US95-03718-2

Query Match 96.9%; Score 5102; DB 5; Length 1000;  
Best Local Similarity 96.7%; Pred. No. 0;  
Matches 968; Conservative 6; Mismatches 17; Indels 10; Gaps 2;

QY 1 MRLAQRDRRLLLVLSWMILETTNODLPVVKCVLISHENNGSAGKPSRYMVRGS 60  
DB 1 MRLAQRDRRLLLVLSWMILETTNODLPVVKCVLISHENNGSAGKPSRYMVRGS 60  
QY 61 PEDLOCTPRQSEGTVEATVEAASGSITLQVQLATPGDLSCLVWFKHSSLCGCPHFD 120  
DB 61 PEDLOCTPRQSEGTVEATVEAASGSITLQVQLATPGDLSCLVWFKHSSLCGCPHFD 120  
QY 121 LQNRGIVSMALLNVTTQAGEYLLHIQSEANVTYVLTNNVDTQYVLRPRYFRKXENQ 180  
DB 121 LQNRGIVSMALLNVTTQAGEYLLHIQSEANVTYVLTNNVDTQYVLRPRYFRKXENQ 180  
QY 181 DALLCISEGVPEPTVEWVLCSSHRESCKEKGPAVVRKEEVLHELFGTDIRCCARNALGR 240  
DB 181 DALLCISEGVPEPTVEWVLCSSHRESCKEKGPAVVRKEEVLHELFGTDIRCCARNALGR 240  
QY 241 ECTKLFITDLNQAQSTLPQLFKVGEPLWIRCKAIHVNHGFLGTWELEDKALEGSYFE 300  
DB 241 ECTKLFITDLNQAQSTLPQLFKVGEPLWIRCKAIHVNHGFLGTWELEDKALEGSYFE 300  
QY 301 MSTYSTNRTMIRILLAFVSSVGRNDGYTTCSSSKHPSQSALVTILEKGINATSSQEEY 360  
DB 301 MSTYSTNRTMIRILLAFVSSVGRNDGYTTCSSSKHPSQSALVTILEKGINATSSQEEY 360  
QY 361 EIDPYEKFCFVRKAYPRICRTWIFSQASFPCEQRLGDEGYSIKFCDHKNKPGYIFY 420  
DB 361 EIDPYEKFCFVRKAYPRICRTWIFSQASFPCEQRLGDEGYSIKFCDHKNKPGYIFY 420  
QY 421 AENDDAQFTKMTLNIRKPKQVLANASASQSCSSDQYPLPSWTWKKCDKSNCTEETIP 480  
DB 421 AENDDAQFTKMTLNIRKPKQVLANASASQSCSSDQYPLPSWTWKKCDKSNCTEETIP 480  
QY 481 EGVWNNKANRVFGQWVSSTLNMSEAGKGLLVKCCAYNSMGTSCTETIFLNSGPPFFIQ 540

DB 481 EGVWNNKANRVFGQWVSSTLNMSEAGKGLLVKCCAYNSMGTSCTETIFLNSGPPFFIQ 540  
QY 541 DNISFYATIGLCLPFIIVLVILVILCHYKKQFRYESOLOMIQVTPGLDNEYFYVDFRDYEF 600  
DB 541 DNISFYATIGLCLPFIIVLVILVILCHYKKQFRYESOLOMIQVTPGLDNEYFYVDFRDYEF 600  
QY 601 DLKWEFFPRENLEFGKVLGSGAFGRVWNAATAYGISTGVSIGVAVKMLKEKADSCKEALM 660  
DB 601 DLKWEFFPRENLEFGKVLGSGAFGRVWNAATAYGISTGVSIGVAVKMLKEKADSCKEALM 660  
QY 661 SELKMWTHLGHHDNIVNLGACTLSGPVYLIFEYCCYGDLLANVLRSKRKFHRTWTIEFK 720  
DB 661 SELKMWTHLGHHDNIVNLGACTLSGPVYLIFEYCCYGDLLANVLRSKRKFHRTWTIEFK 720  
QY 721 EHNFSYPTFOAHSNSMPGSRREVQLHPDLDQSGFNNGSIHSEDEIEYENQRLAESEE 780  
DB 721 EHNFSYPTFOAHSNSMPGSRREVQLHPDLDQSGFNNGSIHSEDEIEYENQRLAESEE 780  
QY 781 EDLNVLTPEDLFCFAYQVAKGMFELEFKSCVHRDLAARNVLTGKVVKICDPGLARDIL 840  
DB 781 EDLNVLTPEDLFCFAYQVAKGMFELEFKSCVHRDLAARNVLTGKVVKICDPGLARDIL 840  
QY 841 SDSSYVVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIFSLGVNYPYFGIPVDAN 900  
DB 841 SDSSYVVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIFSLGVNYPYFGIPVDAN 900  
QY 901 FYKLIQSGFKMEQPPFYATEGIVFVWQSCWAFPSRKRPSPNLTSLPGCOLAEAEACIRT 960  
DB 901 FYKLIQSGFKMEQPPFYATEGIVFVWQSCWAFPSRKRPSPNLTSLPGCOLAEAEACIRT 960  
QY 961 -----SIHLPKQAAPQORGLRAQSPQVRKIHRRS 992  
DB 961 MGVNVPHEPSIYQNRRLPSREAGS-EPPSPQAQVKIHRRS 1000

RESULT 14  
US-08-222-299-4  
; Sequence 4, Application US/08222299  
; Patent No. 5635388  
; GENERAL INFORMATION:  
; APPLICANT: Bennett, Brian D.  
; APPLICANT: Broz, Susan D.  
; APPLICANT: Matthews, William  
; APPLICANT: Zeigler, Francis C.  
; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES T  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/222,299  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hasak, Janet E.  
; REGISTRATION NUMBER: 28,616  
; REFERENCE/DOCKET NUMBER: 879  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-1896  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168



INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 993 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 US-08-222-299-4

Query Match 84.3%; Score 4436.5; DB 1; Length 993;  
 Best Local Similarity 84.0%; Pred. No. 0;  
 Matches 836; Conservative 59; Mismatches 87; Indels 13; Gaps 4;

QY 1 MRALQSRDRRLLLVLSVMIETVTDLPVVKVLSHENNGSSACKPSYRVRGS 60  
 DB 1 MPALA-RDGGQLPLLVVFSAMIFGTITQDLPVVKVLSHENNGSSACKPSYRVRGS 59

QY 61 PEDLOCTPRROSEGTVEAATVEAEGSITLQVLAATPGDLSCLVFKHSSILGCOHPD 120  
 DB 60 PEDLGALRQSSGTVYEAADVDSASITLQVLVDPAGNISCLVFKHSSILGCOHPD 119

QY 121 LQNRGIVSMALNVTTQAGEYLLHIQSERANYTVLFTVNVVDTQLYVLRPFYRQENQ 180  
 DB 120 LQNRGVSMVILKMTQAGEYLLFIQSEATNYTLFTVSVIRNTLLYLRPFYRQENQ 179

QY 181 DALLCISEGVPEPTVEVWVLCSSHRESCKECPAVVRKEEKVHLFGTDIRCCARNALGR 240  
 DB 180 DALVCISESVEPIVEVWVLCDSQESCKEESPAVVKKEEKVHLFGMDIRCCARNELGR 239

QY 241 ECTKLFTIDLNOAQPSTLPQLFLKVGEPWIRCKAIHVNHGFGLTWELEDKALEEGSYFE 300  
 DB 240 ECTRLFTIDLNOPTTLFQLFLKVGEPWIRCKAVHVNHGFGLTWELENKALEEGSYFE 299

QY 301 MSTYSTNRTMIRILLAFVSVGRNDGYTTCSSSKHPSQSALVTILEKGFINATSSOEY 360  
 DB 300 MSTYSTNRTMIRILLAFVSVGRNDGYTTCSSSKHPSQSALVTILEKGFINATSSOEY 359

QY 361 EIDPVEKFCFVFKAYPRICRTWISQASFPCEQRLGEGYISKPCDHKNKPGYIFY 420  
 DB 360 EIDQVEEFCFVFKAYPQICRTWTFSRKSFPCEQRLGEGYISKPCDHKNKPGYIFY 419

QY 421 AENDDAQFTKMTLNRKPKQVLNANASQASCSGDPYLPSTWTKKCSDKSPNCTEETP 480  
 DB 420 AENDDAQFTKMTLNRKPKQVLNANASQASCSGDPYLPSTWTKKCSDKSPNCTEETP 479

QY 481 EGVMNKNKRVFGQWVSSTLMSBAGKLVKCCAYNSMGTSCTEILNPGPPFFIQ 540  
 DB 480 EGVMNKNKRVFGQWVSSTLMSBAGKLVKCCAYNSMGTSCTEILNPGPPFFIQ 539

QY 541 DNISFYATIGLCLPIVILVILCHYKQFYESQLOMIQVTPDLNDFYFYDREY 600  
 DB 540 DNISFYATIGLCLPIVILVILCHYKQFYESQLOMIQVTPDLNDFYFYDREY 599

QY 601 DLKWEFFRENLEFGVLGSGAGRVNNTAYGISKTGVSIOAVVMKKEADSCKEALM 660  
 DB 600 DLKWEFFRENLEFGVLGSGAGRVNNTAYGISKTGVSIOAVVMKKEADSCKEALM 659

QY 661 SELKMTLGHHDNINVLGACTLSPGVYLIPEYCYGDLNLYRSKKEPHRTWTETPK 720  
 DB 660 SELKMTLGHHDNINVLGACTLSPGVYLIPEYCYGDLNLYRSKKEPHRTWTETPK 719

QY 721 EHNFSYTFQAHNSMPSGREGVLPPLDQSGFNNGSIHSEDEIYENOKRLAEDEE 780  
 DB 720 EHNFSYTFQAHNSMPSGREGVLPPLDQSGFNNGSIHSEDEIYENOKRLAEDEE 777

QY 781 EDNLVLTPEDLCLCFAYQVAKGMEFLPKSCVRDLAARNVLTGKVKICDFGLARDIL 840  
 DB 778 EDNLVLTPEDLCLCFAYQVAKGMEFLPKSCVRDLAARNVLTGKVKICDFGLARDIL 837

QY 841 SDSYVVRGNARLPKVMAPESLPFEGYITIKSDVMSYGILLWEIFSLGWNYPGIPVDAN 900  
 DB 838 SDSYVVRGNARLPKVMAPESLPFEGYITIKSDVMSYGILLWEIFSLGWNYPGIPVDAN 897

QY 901 FYKLIQSGFKMEQPFYATGFIYVMSQCAWFDSPFNLTSPGLCOLAEAEACIRT 960  
 DB 900 ECTKLFTIDLNOAQPSTLPQLFLKVGEPWIRCKAIHVNHGFGLTWELEDKALEEGSYFE 300

DB 898 FYKLIQSGFKMEQPFYATGFIYVMSQCAWFDSPFNLTSPGLCOLAEAEAMYQ- 956  
 QY 961 SIHLPKQAP-----QORGGIURAOSPOQVK 986  
 DB 957 NVDGVSCEPHTYQNRPPFSREMDLGLLSPQAQVE 991

RESULT 15  
 US-08-434-878-4  
 ; Sequence 4, Application US/08434878  
 ; Patent No. 5997865  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bennett, Brian D.  
 ; APPLICANT: Broz, Susan D.  
 ; APPLICANT: Matthews, William  
 ; APPLICANT: Zeigler, Francis C.  
 ; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES  
 ; NUMBER OF SEQUENCES: 4  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Genentech, Inc.  
 ; STREET: 460 Point San Bruno Blvd  
 ; CITY: South San Francisco  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 94080  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: patin (Genentech)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/434,878  
 ; FILING DATE:  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Haack, Janet E.  
 ; REGISTRATION NUMBER: 28,616  
 ; REFERENCE/DOCKET NUMBER: 879  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 415/225-1896  
 ; TELEFAX: 415/952-9881  
 ; TELEX: 910/371-7168  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 993 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 US-08-434-878-4

Query Match 84.3%; Score 4436.5; DB 2; Length 993;  
 Best Local Similarity 84.0%; Pred. No. 0;  
 Matches 836; Conservative 59; Mismatches 87; Indels 13; Gaps 4;

QY 1 MRALQSRDRRLLLVLSVMIETVTDLPVVKVLSHENNGSSACKPSYRVRGS 60  
 DB 1 MPALA-RDGGQLPLLVVFSAMIFGTITQDLPVVKVLSHENNGSSACKPSYRVRGS 59

QY 61 PEDLOCTPRROSEGTVEAATVEAEGSITLQVLAATPGDLSCLVFKHSSILGCOHPD 120  
 DB 60 PEDLGALRQSSGTVYEAADVDSASITLQVLVDPAGNISCLVFKHSSILGCOHPD 119

QY 121 LQNRGIVSMALNVTTQAGEYLLHIQSERANYTVLFTVNVVDTQLYVLRPFYRQENQ 180  
 DB 120 LQNRGVSMVILKMTQAGEYLLFIQSEATNYTLFTVSVIRNTLLYLRPFYRQENQ 179

QY 181 DALLCISEGVPEPTVEVWVLCSSHRESCKECPAVVRKEEKVHLFGTDIRCCARNALGR 240  
 DB 180 DALVCISESVEPIVEVWVLCDSQESCKEESPAVVKKEEKVHLFGMDIRCCARNELGR 239

QY 241 ECTKLFTIDLNOAQPSTLPQLFLKVGEPWIRCKAIHVNHGFGLTWELEDKALEEGSYFE 300

```
Db      240  ECTRLFTIDINOTPOITLPOFLKVGEPILWIRKAVHNGFGLTWELNKALEEGNYFE 299
Qy      301  MSTYSNTRMIRILLAFVSSVGNDRYCYTCSSKHPQSALVTILEKGFINATSSOEY 360
Db      300  MSTYSNTRMIRILLAFVSSVANDGYTCSSKHPQSALVTIVEKGFINATNSSEY 359
Qy      361  EIDPYEKFCSVRFKAYPRIRCTWIIFSQAFCPEQORGLDGYISKFCOHKKNKPGYIFY 420
Db      360  EIDYEEFCFSVRFKAYPQIRCTWTSRKFSFCEQKGLDNGYSISKFNHKGHPGEYIFH 419
Qy      421  AENDDAQFTMTNTRKPKQVLANASASQSSDGYPLPSWTWKCKSDKSPNCTEETIP 480
Db      420  AENDDAQFTMTNTRKPKQVLANASASQSCFSDGYPLPSWTWKCKSDKSPNCTEET 479
Qy      481  EGVWNNKANKYFGOWVSSSTLNNMAGKLLVKCCAYNSMGTSCTIFLNSPGPPFIQ 540
Db      480  EGVWNNKANKYFGOWVSSSTLNNMAGKLLVKCCAYNSMGTSCTIFLNSPGPPFIQ 539
Qy      541  DNISFYATIGLCPLFIWLVILVILCHYKKQFRYESQLOMIQVTGPLDNEYFYVDFRDY 600
Db      540  DNISFYATIGVCLLFIWLVILVILCHYKKQFRYESQLOMIQVTGSSDNEYFYVDFREY 599
Qy      601  DLKWEPPRENLEFGKVLGSAFGKQVNNATAYGISKTVSIQVAVKMLKEKADSCKEALM 660
Db      600  DLKWEPPRENLEFGKVLGSAFGKQVNNATAYGISKTVSIQVAVKMLKEKADSSREALM 659
Qy      661  SELKMMTHLGHDNIYNLLGACTLSGPVYLIPEYCCYGDLLNLYRSKREKPHRTWTEIF 720
Db      660  SELKMMTHLGSHENIYNLLGACTLSGPYLIPEYCCYGDLLNLYRSKREKPHRTWTEIF 719
Qy      721  EHNFSYPTFOAHSNSMPSREVQLHPDQLSGFNNGNSIHSEDEIEYENOKRLAEBEE 780
Db      720  EHNFSYPTFOAHSNSMPSREVQLHPDQLSGFNNGNSIHSEDEIEYENOKRLAEBEE 777
Qy      781  EDNLVLTFFEDLLCFAYQVAKMEFLBFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
Db      778  EDNLVLTFFEDLLCFAYQVAKMEFLBFKSCVHRDLAARNVLVTHGKVKICDFGLARDIM 837
Qy      841  SDSSVYVRGNARLPVKWMAPESLFEGYITIKSDVNSYGILLWEIFSLGVNPPYGIPIVDAN 900
Db      838  SDSSVYVRGNARLPVKWMAPESLFEGYITIKSDVNSYGILLWEIFSLGVNPPYGIPIVDAN 897
Qy      901  FYKLIQSGFKMEOPFYATEGIYFVMSQWAFDSRKRPSPNLTSLGCOLAEBEACIRT 960
Db      898  FYKLIQSGFKMEOPFYATEGIYFVMSQWAFDSRKRPSPNLTSLGCOLAEBEACIRT 956
Qy      961  SIHLPKQAAP-----QQRGLRAQSPQVVK 986
Db      957  NVDGPVSECPHTYQNRPPFSREMDLGLLSPQAQVE 991
```

Search completed: August 26, 2003, 07:29:51  
Job time : 26.4892 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 14:24:01 ; Search time 16319.9 Seconds  
(without alignments)  
8776.084 Million cell updates/sec

Title: US-09-919-408A-3  
Perfect score: 3501  
Sequence: 1 CGAGCGGCATCGAGGGCT.....AAAAAAAAAAAAAAAAAAAA 3501

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues  
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_hcg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sv.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_ov.\*
- 22: em\_pat.\*
- 23: em\_ph.\*
- 24: em\_pi.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_hcg\_hum.\*
- 31: em\_hcg\_inv.\*
- 32: em\_hcg\_other.\*
- 33: em\_hcg\_mus.\*
- 34: em\_hcg\_pln.\*
- 35: em\_hcg\_rod.\*
- 36: em\_hcg\_mam.\*
- 37: em\_hcg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_hgo\_hum.\*
- 40: em\_hgo\_mus.\*
- 41: em\_hgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	3501	100.0	3501	6	AR005212 Sequence
2	3501	100.0	3501	6	AR071704 Sequence
3	3501	100.0	3501	6	125170 Sequence 3
4	3501	100.0	3501	6	140601 Sequence 3
5	3475	99.3	3475	6	AX695522 Sequence
6	3475	99.3	3475	6	AX695522 Sequence 3
7	3475	99.3	3475	6	144733 Sequence 1
8	3475	99.3	3475	9	HSU02687
9	3475	99.3	3475	6	140169 Sequence 1
10	3475	99.3	3475	6	140169 Sequence 1
11	3475	99.3	3475	6	140169 Sequence 1
12	3475	99.3	3475	6	140169 Sequence 1
13	3475	99.3	3475	6	140169 Sequence 1
14	3475	99.3	3475	6	140169 Sequence 1
15	3475	99.3	3475	6	140169 Sequence 1
16	3475	99.3	3475	6	140169 Sequence 1
17	3475	99.3	3475	6	140169 Sequence 1
18	3475	99.3	3475	6	140169 Sequence 1
19	3475	99.3	3475	6	140169 Sequence 1
20	3475	99.3	3475	6	140169 Sequence 1
21	3475	99.3	3475	6	140169 Sequence 1
22	3475	99.3	3475	6	140169 Sequence 1
23	3475	99.3	3475	6	140169 Sequence 1
24	3475	99.3	3475	6	140169 Sequence 1
25	3475	99.3	3475	6	140169 Sequence 1
26	3475	99.3	3475	6	140169 Sequence 1
27	3475	99.3	3475	6	140169 Sequence 1
28	3475	99.3	3475	6	140169 Sequence 1
29	3475	99.3	3475	6	140169 Sequence 1
30	3475	99.3	3475	6	140169 Sequence 1
31	3475	99.3	3475	6	140169 Sequence 1
32	3475	99.3	3475	6	140169 Sequence 1
33	3475	99.3	3475	6	140169 Sequence 1
34	3475	99.3	3475	6	140169 Sequence 1
35	3475	99.3	3475	6	140169 Sequence 1
36	3475	99.3	3475	6	140169 Sequence 1
37	3475	99.3	3475	6	140169 Sequence 1
38	3475	99.3	3475	6	140169 Sequence 1
39	3475	99.3	3475	6	140169 Sequence 1
40	3475	99.3	3475	6	140169 Sequence 1
41	3475	99.3	3475	6	140169 Sequence 1
42	3475	99.3	3475	6	140169 Sequence 1
43	3475	99.3	3475	6	140169 Sequence 1
44	3475	99.3	3475	6	140169 Sequence 1
45	3475	99.3	3475	6	140169 Sequence 1

ALIGNMENTS

RESULT 1	AR005212	Sequence 3	from patent US 5747651.	3501 bp	DNA	linear	PAT 04-DEC-1998
LOCUS	AR005212	Sequence 3	from patent US 5747651.	3501 bp	DNA	linear	PAT 04-DEC-1998
DEFINITION	AR005212	Sequence 3	from patent US 5747651.	3501 bp	DNA	linear	PAT 04-DEC-1998
ACCESSION	AR005212	Sequence 3	from patent US 5747651.	3501 bp	DNA	linear	PAT 04-DEC-1998
VERSION	AR005212.1	GI:3966091					
KEYWORDS	Unknown.						
SOURCE	Unknown.						
ORGANISM	Unknown.						
REFERENCE	1 (bases 1 to 3501)						
AUTHORS	Lemischka, I.R.						
TITLE	Antibodies against tyrosine kinase receptor flk-1						
JOURNAL	Patent: US 5747651-A 3 05-MAY-1998;						
FEATURES	Location/Qualifiers						

[illegible]

2041	Db	CTCAAGATGATGACCCAGCTGGGAGCCAGAAATTGTGAACTATTGTGGGGCGTGC	2100
2101	QY	ACACTGTGAGGACCAATTTACTTTGATTTTGAATACTGTGCTATGGTGATCTTCTCAAC	2160
2101	Db	ACACTGTGAGGACCAATTTACTTTGATTTTGAATACTGTGCTATGGTGATCTTCTCAAC	2160
2161	QY	TATCTAAGAGTAAGAAAGAGAAAAATTTCAAGACACTTGGACAGAGATTTTCAAGGAACAC	2220
2161	Db	TATCTAAGAGTAAGAAAGAGAAAAATTTCAAGACACTTGGACAGAGATTTTCAAGGAACAC	2220
2221	QY	AATTTTCAGTTTTTACCCCACTTTTCCAATCATCTCAAATTCAGCATGCTGGTCTCAAGA	2280
2221	Db	AATTTTCAGTTTTTACCCCACTTTTCCAATCATCTCAAATTCAGCATGCTGGTCTCAAGA	2280
2281	QY	GAAGTTCAGATACACCCGGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCCAC	2340
2281	Db	GAAGTTCAGATACACCCGGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCCAC	2340
2341	QY	TCCTGAAGATCAAAATGGAATATGAAAAACAAAAAGGCTGGAAGACAGGAGGACTTTGAAAT	2400
2341	Db	TCCTGAAGATCAAAATGGAATATGAAAAACAAAAAGGCTGGAAGACAGGAGGACTTTGAAAT	2400
2401	QY	GTGCTTACATTTGAAGACTTCTTTGCTTTGCTATATCAAGTTCGCAAGGAATGGAATTT	2460
2401	Db	GTGCTTACATTTGAAGACTTCTTTGCTTTGCTATATCAAGTTCGCAAGGAATGGAATTT	2460
2461	QY	CTGGAAATTTAAGCTGTGTGTTTACAGAGACCTGGCCGCCAGGAACGTGTTGTCACCCAC	2520
2461	Db	CTGGAAATTTAAGCTGTGTGTTTACAGAGACCTGGCCGCCAGGAACGTGTTGTCACCCAC	2520
2521	QY	GGGNAAGTGTGAAGATATGTGACTTTGGATTTGCTCGAGATATCATAGTGAATTCCAAC	2580
2521	Db	GGGNAAGTGTGAAGATATGTGACTTTGGATTTGCTCGAGATATCATAGTGAATTCCAAC	2580
2581	QY	TATGTTCTCAGGGGCAATGCCGCTCTGCCGTGTAATGGAATGAGTGGCCCCGAAAGCCCTGTTT	2640
2581	Db	TATGTTCTCAGGGGCAATGCCGCTCTGCCGTGTAATGGAATGAGTGGCCCCGAAAGCCCTGTTT	2640
2641	QY	GAAGGCATCTACACCAATTAAGATGATGTCTGTGTCATATGGAATATTACTGTGGGAATC	2700
2641	Db	GAAGGCATCTACACCAATTAAGATGATGTCTGTGTCATATGGAATATTACTGTGGGAATC	2700
2701	QY	TTCTCACTTCGGTGAATCCTTACCCCTGGCATTCGGGTTCAGTCTACTTCTACAAACTG	2760
2701	Db	TTCTCACTTCGGTGAATCCTTACCCCTGGCATTCGGGTTCAGTCTACTTCTACAAACTG	2760
2761	QY	ATTCAAATGGATTTAAATGGATCAGCCATTTTATGCTACAGAAGAAATATACATTATA	2820
2761	Db	ATTCAAATGGATTTAAATGGATCAGCCATTTTATGCTACAGAAGAAATATACATTATA	2820
2821	QY	ATGCAATCCTGCTGGGCTTTTGACTCAAGGAACGGCCATCTTCCCTAAATTGACTTCG	2880
2821	Db	ATGCAATCCTGCTGGGCTTTTGACTCAAGGAACGGCCATCTTCCCTAAATTGACTTCG	2880
2881	QY	TTTTTTAGGATCTCAGCTGGCAGATGTCAGAGAACGGATGTATCAGATGTGGATGGCCGT	2940
2881	Db	TTTTTTAGGATCTCAGCTGGCAGATGTCAGAGAACGGATGTATCAGATGTGGATGGCCGT	2940
2941	QY	GTTTTCGGAATCTCCTCACACCTACCAAAAACAGGCGACCTTTTCAGCAGAGATGGATTTG	3000
2941	Db	GTTTTCGGAATCTCCTCACACCTACCAAAAACAGGCGACCTTTTCAGCAGAGATGGATTTG	3000
3001	QY	GGGCTATCTCTCCGCGAGGCTCAGGTGGAATTCGTAGAGGAACAATTTAGTTTAAAGG	3060
3001	Db	GGGCTATCTCTCCGCGAGGCTCAGGTGGAATTCGTAGAGGAACAATTTAGTTTAAAGG	3060
3061	QY	ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTAATTTTCAT	3120
3061	Db	ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTAATTTTCAT	3120
3121	QY	CACCTAAAAGAAAACTATTTATCAACTGCTGCTTCACCAGACTTTTCTCTAGAAAGCCGTCT	3180

Db	3121	CACTAAAGAAAAATCTATTATCAACTGCTGCTTCAACGAGACTTTTCTCTAGAAGCCGTCT	3118
Qy	3181	GCCTTTACTCTTCTTTTCAAAGGGAATTTTGTAAAAATCAAATCATCTGTCAAGAGCAG	3240
Db	3181	GCCTTTACTCTTCTTTTCAAAGGGAATTTTGTAAAAATCAAATCATCTGTCAAGAGCAG	3240
Qy	3241	GAGGAGCTGATAATGAACCTTTATTGGAGCAATGATCTGCAATCCAAGGCCCTTCTCAGGCCG	3300
Db	3241	GAGGAGCTGATAATGAACCTTTATTGGAGCAATGATCTGCAATCCAAGGCCCTTCTCAGGCCG	3300
Qy	3301	GCTTGAGTGAATTTGCTGCTACTGGAAGTACAGTATATCTTGTAAATACATAAAACAAAAGC	3360
Db	3301	GCTTGAGTGAATTTGCTGCTACTGGAAGTACAGTATATCTTGTAAATACATAAAACAAAAGC	3360
Qy	3361	ATTTTGTCAAGGAGAAGCTAATATGATTTTTTAAAGTCTATGTTTTAAAAATAATATGTAAA	3420
Db	3361	ATTTTGTCAAGGAGAAGCTAATATGATTTTTTAAAGTCTATGTTTTAAAAATAATATGTAAA	3420
Qy	3421	TTTTTTCAGCTATTATTAGTGATATATTTTATGGGTGGGAATAAAATTTCTACTACAGAAAAA	3480
Db	3421	TTTTTTCAGCTATTATTAGTGATATATTTTATGGGTGGGAATAAAATTTCTACTACAGAAAAA	3480
Qy	3481	AAAAAAAAAAAAAAAAAAAAA 3501	
Db	3481	AAAAAAAAAAAAAAAAAAAAA 3501	
RESULT 2			
AR071704	AR071704	3501 bp	DNA
LOCUS	AR071704	3501 bp	linear
DEFINITION	Sequence 3 from patent US 5912133.		PAT 18-FEB-2000
ACCESSION	AR071704		
VERSION	AR071704.1	GI:7222592	
KEYWORDS	unknown.		
SOURCE	unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 3501)		
AUTHORS	Lemischka, I.R.		
TITLE	Method for isolating stem cells expressing flk-1 receptors		
JOURNAL	Patent: US 5912133-A 3 15-JUN-1999;		
FEATURES	Location/Qualifiers		
source	1..3501		
BASE COUNT	1068 a 709 c 784 g 940 t		
ORIGIN	/organism="unknown"		
Query Match	100.0%;	Score 3501;	DB 6; Length 3501;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 3501;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	1	CGAGGCGGCATCCGAGGGCTGGGCGGCGCCCTGGGGACCCCGGGCTCCGGAGGCCATG	60
Db	1	CGAGGCGGCATCCGAGGGCTGGGCGGCGCCCTGGGGACCCCGGGCTCCGGAGGCCATG	60
Qy	61	CCGGCGTTGGCGCGCAGCGGGGACCGGTGCGCGCTGCTCGTGTGTTTTTCTGCAATGATA	120
Db	61	CCGGCGTTGGCGCGCAGCGGGGACCGGTGCGCGCTGCTCGTGTGTTTTTCTGCAATGATA	120
Qy	121	TTTGGGACTATTCAAATCAAGATCTGCCCTGTGATCAAGTGTGTTTTTAATCAATCAATAAG	180
Db	121	TTTGGGACTATTCAAATCAAGATCTGCCCTGTGATCAAGTGTGTTTTTAATCAATCAATAAG	180
Qy	181	AACATGATTCATCAGTGGGGAAGTCAATCATATCCCATGGTATCAGAAATCCCGGAA	240
Db	181	AACATGATTCATCAGTGGGGAAGTCAATCATATCCCATGGTATCAGAAATCCCGGAA	240
Qy	241	GACCTCGGGTGTGCTTTGAGACCCAGAGCTCAGGGACAGTGTACGAAGCTGCCGCTGTG	300
Db	241	GACCTCGGGTGTGCTTTGAGACCCAGAGCTCAGGGACAGTGTACGAAGCTGCCGCTGTG	300
Qy	301	GAAGTGATGATCTGCTTCCATCACAATGCAAGTGTGCTCGATGCCCGAGGAAATTT	360



2521 GGAAAGTGGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTATTCAC 2580  
 2521 GGAAAGTGGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTATTCAC 2580  
 2581 TATGTTGTAGGGCAATGCCGCTGCTCCCTGTTAAATGGATGGCCCGCAAGCCTGTTT 2640  
 2581 TATGTTGTAGGGCAATGCCGCTGCTCCCTGTTAAATGGATGGCCCGCAAGCCTGTTT 2640  
 2641 GAAGGATCTACACCATTAAGAGTATGTTGGTTCATATGGAATATTTACTGTGGGAATC 2700  
 2641 GAAGGATCTACACCATTAAGAGTATGTTGGTTCATATGGAATATTTACTGTGGGAATC 2700  
 2701 TTCTCACTTGGTGAATCTTACCTTGGCAATCCCGTTGATGTAATCTTCTACAAACTG 2760  
 2701 TTCTCACTTGGTGAATCTTACCTTGGCAATCCCGTTGATGTAATCTTCTACAAACTG 2760  
 2761 ATTCAAATGGATTTAAATGGATCAAGCAATTTTATGCTACAGAGAAATATATCATATA 2820  
 2761 ATTCAAATGGATTTAAATGGATCAAGCAATTTTATGCTACAGAGAAATATATCATATA 2820  
 2821 ATGCAATCTGCTGGGCTTTTGACTCAAGAAAGGGCCATCTTCCCTAATTTGACTTCG 2880  
 2821 ATGCAATCTGCTGGGCTTTTGACTCAAGAAAGGGCCATCTTCCCTAATTTGACTTCG 2880  
 2881 TTTTATAGGATGTGAGCTGGCAGATGCAAGAGAGCGATGTATCAGATGTGGATGGCCGT 2940  
 2881 TTTTATAGGATGTGAGCTGGCAGATGCAAGAGAGCGATGTATCAGATGTGGATGGCCGT 2940  
 2941 GTTTCGGAATGTCTCAACCTACCAAAACAGGCGACCTTTTACGAGAGAGATGGAATTG 3000  
 2941 GTTTCGGAATGTCTCAACCTACCAAAACAGGCGACCTTTTACGAGAGAGATGGAATTG 3000  
 3001 GGGCTACTCTCTCGCAGGCTCAGGTGCAAGATTCGTTAGAGAAACAATTTAGTTTAAAG 3060  
 3001 GGGCTACTCTCTCGCAGGCTCAGGTGCAAGATTCGTTAGAGAAACAATTTAGTTTAAAG 3060  
 3061 ACTTCACTCCCTCAGCTATCCCTTAAAGAGCTGATGATTTACCAAAACAAGATTTATTTT 3120  
 3061 ACTTCACTCCCTCAGCTATCCCTTAAAGAGCTGATGATTTACCAAAACAAGATTTATTTT 3120  
 3121 CACTAAAGAAATCTATTATCAACTGCTCTTCAAGAGCTTTTCTTAAAGCCGCTCT 3180  
 3121 CACTAAAGAAATCTATTATCAACTGCTCTTCAAGAGCTTTTCTTAAAGCCGCTCT 3180  
 3181 GCGTTTACTCTGTTTCAAGGAGCTTTTGTAAATCAATCATCTGTCACAAAGCAG 3240  
 3181 GCGTTTACTCTGTTTCAAGGAGCTTTTGTAAATCAATCATCTGTCACAAAGCAG 3240  
 3241 GAGGAGCTGATATGAATTTATTTGAGGAGCTTGAATCTGCAATCCAGGCTTCTCAGGCG 3300  
 3241 GAGGAGCTGATATGAATTTATTTGAGGAGCTTGAATCTGCAATCCAGGCTTCTCAGGCG 3300  
 3301 GCTTGAAGTAAATGTTGATCTGAAGTACAGTATATCTTGTAAATACATAAAACAAAGC 3360  
 3301 GCTTGAAGTAAATGTTGATCTGAAGTACAGTATATCTTGTAAATACATAAAACAAAGC 3360  
 3361 ATTTCGTAAGGAGAGCTAATATGATTTTAAAGTCTATGTTTAAATATATATGTA 3420  
 3361 ATTTCGTAAGGAGAGCTAATATGATTTTAAAGTCTATGTTTAAATATATATGTA 3420  
 3421 TTTTTCAGCTATTTAGTATATATTTATGGTGGGAAATATAATTTCTACTACAGAAAA 3480  
 3421 TTTTTCAGCTATTTAGTATATATTTATGGTGGGAAATATAATTTCTACTACAGAAAA 3480  
 3481 AAAAAAAAAAAAAAAAAAAAAA 3501  
 3481 AAAAAAAAAAAAAAAAAAAAAA 3501

RESULT 3  
 125170  
 LOCUS  
 DEFINITION Sequence 3 from patent US 5548065.  
 125170 3501 bp DNA linear PAT 07-OCT-1996

125170  
 125170.1 GI:1605040  
 Unknown.  
 SOURCE  
 ORGANISM  
 Unclassified.  
 REFERENCE 1 (bases 1 to 3501)  
 Lemischka, I.R.  
 Tyrosine kinase receptor human flk-2-specific antibodies  
 Patent: US 5548065-A 3 20-AUG-1996;  
 Location/Qualifiers  
 FEATURES  
 1..3501  
 source  
 BASE COUNT 1068 a 709 c 784 g 940 t  
 ORIGIN  
 Query Match 100.0%; Score 3501; DB 6; Length 3501;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGAGGCGGCATCCGAGGGCTGGGCGGCCCTTGGGGGACCCCGGGCTCCGGAGGCCATG 60  
 Db 1 CGAGGCGGCATCCGAGGGCTGGGCGGCCCTTGGGGGACCCCGGGCTCCGGAGGCCATG 60  
 QY 61 CCGGCGTTGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120  
 Db 61 CCGGCGTTGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120  
 QY 121 TTTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATCAATAAG 180  
 Db 121 TTTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATCAATAAG 180  
 QY 181 AACAAATGATTCATCAGTGGGAAAGTCAATCAATATCCCATGGTATCAGAAATCCCGGAA 240  
 Db 181 AACAAATGATTCATCAGTGGGAAAGTCAATCAATATCCCATGGTATCAGAAATCCCGGAA 240  
 QY 241 GACCTGGGTGTCGGTTGAGACCCGAGCTCAGGAGCAGTGTACGAGCTGCCGCTGTG 300  
 Db 241 GACCTGGGTGTCGGTTGAGACCCGAGCTCAGGAGCAGTGTACGAGCTGCCGCTGTG 300  
 QY 301 GAAGTGGATGTATCTGCTTCCATCAGCTGCAAGTGTGTCGATGCCCGGAAACATT 360  
 Db 301 GAAGTGGATGTATCTGCTTCCATCAGCTGCAAGTGTGTCGATGCCCGGAAACATT 360  
 QY 361 TCCTGTCTCTGGGTCTTTAAGCAGCTCCCTGAAATGCCAGCCACATTTTGAATTTACAA 420  
 Db 361 TCCTGTCTCTGGGTCTTTAAGCAGCTCCCTGAAATGCCAGCCACATTTTGAATTTACAA 420  
 QY 421 AACAGAGGAGTGTGTTCCATGGTCAATTTGAAATGACAGAAACCCAGCTGGAGATAC 480  
 Db 421 AACAGAGGAGTGTGTTCCATGGTCAATTTGAAATGACAGAAACCCAGCTGGAGATAC 480  
 QY 481 CTACTTTTATTCAGAGTGAAGCTACCAATTCACATATTTTACAGTCAAGTATTAAGA 540  
 Db 481 CTACTTTTATTCAGAGTGAAGCTACCAATTCACATATTTTACAGTCAAGTATTAAGA 540  
 QY 541 AATACCTCTGTTTACACATTTAAGAGACCTTACTTTAGAAAAATGGAAAAACAGAGCC 600  
 Db 541 AATACCTCTGTTTACACATTTAAGAGACCTTACTTTAGAAAAATGGAAAAACAGAGCC 600  
 QY 601 CTGGTCTGCATATCTCAGAGCGTTCCAGAGCCGATCGTGAATGGGTGCTTTTGCATTC 660  
 Db 601 CTGGTCTGCATATCTCAGAGCGTTCCAGAGCCGATCGTGAATGGGTGCTTTTGCATTC 660  
 QY 661 CAGGGGGAAGCTGTAAAGAGAAAGTCCAGCTGTTTAAAGAGGAAAGAAAGTCTT 720  
 Db 661 CAGGGGGAAGCTGTAAAGAGAAAGTCCAGCTGTTTAAAGAGGAAAGAAAGTCTT 720  
 QY 721 CATCAATTTATTTGGGACGGACATTAAGTGTCTGTGCCAGAAATGAATCGGGCAGGAAATGC 780  
 Db 721 CATCAATTTATTTGGGACGGACATTAAGTGTCTGTGCCAGAAATGAATCGGGCAGGAAATGC 780  
 QY 781 ACCAGGCTGTTCACAAATAGATCTAAATCAAACTCCTCAGACCACATTTGCCCAATTTT 840



Db 781 ACCAGGCTGTTCAAAATAGATCTAAATCAAACTCCTCAGACCAATATGCCCAATATATT 840  
Qy CTTAAAGTAGGGAAACCTTATGATAGAGTGCAAAAGCTGTTCAATGTGAACCATGGATTC 900  
Db CTTAAAGTAGGGAAACCTTATGATAGAGTGCAAAAGCTGTTCAATGTGAACCATGGATTC 900  
Qy GGGCTACCTGGGAATTTAGAAAAAAGAGCACTCAGAGGGGCAACTACTTTTGAGATGAT 960  
Db GGGCTACCTGGGAATTTAGAAAAAAGAGCACTCAGAGGGGCAACTACTTTTGAGATGAT 960  
Qy ACCTATTCAACAAACAGACATATGATACGGATTCGTGTTGCTTTGTTGATCATCATGTCGCA 1020  
Db ACCTATTCAACAAACAGACATATGATACGGATTCGTGTTGCTTTGTTGATCATCATGTCGCA 1020  
Qy AGAAACGACACCGGATCTACACTGTTGTTCTCTCTCAAGCATCCAGTCAATCAGCTTTG 1080  
Db AGAAACGACACCGGATCTACACTGTTGTTCTCTCTCAAGCATCCAGTCAATCAGCTTTG 1080  
Qy GTTACCATCTGAGAAAGGGATTTATAAATGCTACCAATTCAGAGTGAAGATTTATGAATTT 1140  
Db GTTACCATCTGAGAAAGGGATTTATAAATGCTACCAATTCAGAGTGAAGATTTATGAATTT 1140  
Qy GACCAATATGAAAGATTTGTTTCTCTGTCAGGTTTAAAGCCTACCCACAAATCAGATGT 1200  
Db GACCAATATGAAAGATTTGTTTCTCTGTCAGGTTTAAAGCCTACCCACAAATCAGATGT 1200  
Qy ACGTGGACCTTCTCGAAATCATTTCTCTGTCAGCAAAAGGCTCTTGATAACGGATAC 1260  
Db ACGTGGACCTTCTCGAAATCATTTCTCTGTCAGCAAAAGGCTCTTGATAACGGATAC 1260  
Qy AGCATATCCAAAGTTTTCATATCAAGCACAGCCAGGAGGAATATATATCCATGCGAAA 1320  
Db AGCATATCCAAAGTTTTCATATCAAGCACAGCCAGGAGGAATATATATCCATGCGAAA 1320  
Qy AATGATGATGCCAAATTTACCAAAATGTTACGCTGTAATATAAGAAAGAAACCTCAAGTG 1380  
Db AATGATGATGCCAAATTTACCAAAATGTTACGCTGTAATATAAGAAAGAAACCTCAAGTG 1380  
Qy CTGCGAAGCATCGGCAAGTCTCAGCAAGTCTGCTGTTCTCGATGATACCCATACCATCT 1440  
Db CTGCGAAGCATCGGCAAGTCTCAGCAAGTCTGCTGTTCTCGATGATACCCATACCATCT 1440  
Qy TGGAACCTGGAAAGATGTTTCAGCAAGTCTCCCACTGCAAGAGAGATCACAGAAGGA 1500  
Db TGGAACCTGGAAAGATGTTTCAGCAAGTCTCCCACTGCAAGAGAGATCACAGAAGGA 1500  
Qy GTCTGGAATAGAAAGGCTTACAGAAAGTGTTCGACAGTGGGTGTCGAGCAGTACTCTA 1560  
Db GTCTGGAATAGAAAGGCTTACAGAAAGTGTTCGACAGTGGGTGTCGAGCAGTACTCTA 1560  
Qy AACATGATGAAAGGCTTACAAAGGCTTCTGCTCAAGTGTGTCATACAAATTCCTTTGGC 1620  
Db AACATGATGAAAGGCTTACAAAGGCTTCTGCTCAAGTGTGTCATACAAATTCCTTTGGC 1620  
Qy ACATCTTTGTGAGAGCATCTTTTAAACTCTCCAGGCCCCCTTCCCTTTCATCCAGAACA 1680  
Db ACATCTTTGTGAGAGCATCTTTTAAACTCTCCAGGCCCCCTTCCCTTTCATCCAGAACA 1680  
Qy ATCTATCTATGCAACAAATGGTGTGTTGCTCTCTTCAATGTCGTTTAAACCTGCTA 1740  
Db ATCTATCTATGCAACAAATGGTGTGTTGCTCTCTTCAATGTCGTTTAAACCTGCTA 1740  
Qy ATTTGTACAAAGTACAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGTG 1800  
Db ATTTGTACAAAGTACAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGTG 1800  
Qy ACCGGTCTCCTCAGATATGATGATCTCTACGTTGATTTTCAGAGAAATATGAATATGATCTC 1860  
Db ACCGGTCTCCTCAGATATGATGATCTCTACGTTGATTTTCAGAGAAATATGAATATGATCTC 1860  
Qy AAATGGAGTTTCCAGAGAAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920  
Db AAATGGAGTTTCCAGAGAAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT

Db 1861 AAATGGAGTTTCCAGAGAAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920  
Qy GGAAGAGTATGAACGCAACAGCTTATGGAATTTAGCAAAAACAGAGTCTCAATCCAGGTT 1980  
Db GGAAGAGTATGAACGCAACAGCTTATGGAATTTAGCAAAAACAGAGTCTCAATCCAGGTT 1980  
Qy GCGTCAAAATGCTGAAAGAAAAAGAGCAGCTCTGAAAGAGAGGCACTCATGTTCAGAA 2040  
Db GCGTCAAAATGCTGAAAGAAAAAGAGCAGCTCTGAAAGAGAGGCACTCATGTTCAGAA 2040  
Qy CTCAAGATGATGACCCAGCTGGGAAGCCACAGAAATATTTGGAACCTGCTGGGGGCGTGC 2100  
Db CTCAAGATGATGACCCAGCTGGGAAGCCACAGAAATATTTGGAACCTGCTGGGGGCGTGC 2100  
Qy ACCTGTGAGGACCAATTTTACTGATTTTGAATCTGTTGCTATGTTGATCTTCTCAAC 2160  
Db ACCTGTGAGGACCAATTTTACTGATTTTGAATCTGTTGCTATGTTGATCTTCTCAAC 2160  
Qy TATCTAAGAGTAAAGAGAAAAATTTTCAGAGCTTGGACAGAGATTTTTCAGAGAACAC 2220  
Db TATCTAAGAGTAAAGAGAAAAATTTTCAGAGCTTGGACAGAGATTTTTCAGAGAACAC 2220  
Qy AATTTGAGTTTACCCCACTTTCCAAATCAGATCCAAATTTCCAGCATGCTGTTCAAGA 2280  
Db AATTTGAGTTTACCCCACTTTCCAAATCAGATCCAAATTTCCAGCATGCTGTTCAAGA 2280  
Qy GAAGTTCAGATACACCCGGACTCGGATCAAACTCTCAGGGCTTCAATGGAAATTCATTTCC 2340  
Db GAAGTTCAGATACACCCGGACTCGGATCAAACTCTCAGGGCTTCAATGGAAATTCATTTCC 2340  
Qy TCTGAAGATCAAAATGTAATATGAACCAAAAAAGGCTGGAAGAGGAGGACCTTGAAT 2400  
Db TCTGAAGATCAAAATGTAATATGAACCAAAAAAGGCTGGAAGAGGAGGACCTTGAAT 2400  
Qy GTGCTTACATTTGAAGATCTTTGCTTGTGCTATCAAGTTGCCAAGAGAAATGGAATTT 2460  
Db GTGCTTACATTTGAAGATCTTTGCTTGTGCTATCAAGTTGCCAAGAGAAATGGAATTT 2460  
Qy CTGGAATTTAAGTCTGTTTCAGAGACCTGCGCCGCGAGGAACGCTTGTCAACCAC 2520  
Db CTGGAATTTAAGTCTGTTTCAGAGACCTGCGCCGCGAGGAACGCTTGTCAACCAC 2520  
Qy GGAAGAGTGTGAAGATATGATGATTTGGATTTGCTCGAGATATCATGAGTGAATCCAAAC 2580  
Db GGAAGAGTGTGAAGATATGATGATTTGGATTTGCTCGAGATATCATGAGTGAATCCAAAC 2580  
Qy TATGTTGTGAGGGCAATGCGGCTGCTGCTGTAATAATGGATGGCCCGGAAAGCCTGTTT 2640  
Db TATGTTGTGAGGGCAATGCGGCTGCTGCTGTAATAATGGATGGCCCGGAAAGCCTGTTT 2640  
Qy GAAGGCATCTACACCATTAAGAGTGTCTGTTGATGATGAAATTTACTTGTGGAAATC 2700  
Db GAAGGCATCTACACCATTAAGAGTGTCTGTTGATGATGAAATTTACTTGTGGAAATC 2700  
Qy TTCTCACTTGTGTAATCTTACCTTGGCAATTTCCGGTTGATGCTAACTTTCACAAACCTG 2760  
Db TTCTCACTTGTGTAATCTTACCTTGGCAATTTCCGGTTGATGCTAACTTTCACAAACCTG 2760  
Qy ATTCAAAATGGATTTAAATGGATTCAGCCATTTTATGCTACAGAGAAATATACATTATA 2820  
Db ATTCAAAATGGATTTAAATGGATTCAGCCATTTTATGCTACAGAGAAATATACATTATA 2820  
Qy ATCCAAATCTGCTGGGCTTTTGTACTCAAGAAACGGCCATCTTCCCTTAATTTGACTTCG 2880  
Db ATCCAAATCTGCTGGGCTTTTGTACTCAAGAAACGGCCATCTTCCCTTAATTTGACTTCG 2880  
Qy TTTTATAGGATGTGAGCTGGCAGATGCAAGAAAGAGCGATGTATCAGATGTGGATGGCGGT 2940  
Db TTTTATAGGATGTGAGCTGGCAGATGCAAGAAAGAGCGATGTATCAGATGTGGATGGCGGT 2940  
Qy GTTTCGGAATGCTCCTCACACCTACCAAAAACAGGCACTTTTCAGACAGAGATCGATTTG 3000  
Db GTTTCGGAATGCTCCTCACACCTACCAAAAACAGGCACTTTTCAGACAGAGATCGATTTG 3000

QY	3001	GGGCTACTCTCTCCGAGGCTCAGGTGCGAGTTCGTAGAGGAAACAATTTAGTTTAAAGG	3060
Db	3001	GGGCTACTCTCTCCGAGGCTCAGGTGCGAGTTCGTAGAGGAAACAATTTAGTTTAAAGG	3060
QY	3061	ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTAAATTTTCAT	3120
Db	3061	ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTAAATTTTCAT	3120
QY	3121	CACATAAAGAAATCTATATCACTGCTTCCACGAGACTTTTCTCTAGAGGCCGTCT	3180
Db	3121	CACATAAAGAAATCTATATCACTGCTTCCACGAGACTTTTCTCTAGAGGCCGTCT	3180
QY	3181	CGGTTTACTCTGTTTTCACAGGAGCTTTTGTAAATCAAAATCCTGTCTACAGGCAG	3240
Db	3181	CGGTTTACTCTGTTTTCACAGGAGCTTTTGTAAATCAAAATCCTGTCTACAGGCAG	3240
QY	3241	GAGGAGCTGATTAATGAACCTTTATTGGAGCAATGATCTGCATCCAGGCCCTTCAGGCCG	3300
Db	3241	GAGGAGCTGATTAATGAACCTTTATTGGAGCAATGATCTGCATCCAGGCCCTTCAGGCCG	3300
QY	3301	GCTTGAGTGAATTTGTGACCTGGAAGTACAGTATATCTTGTAAATACATAAAAAACAAGC	3360
Db	3301	GCTTGAGTGAATTTGTGACCTGGAAGTACAGTATATCTTGTAAATACATAAAAAACAAGC	3360
QY	3361	ATTTGCTAGGAGCACTAATATGATTTTTTAAGTCTATGTTTAAATAAATATGTAAA	3420
Db	3361	ATTTGCTAGGAGCACTAATATGATTTTTTAAGTCTATGTTTAAATAAATATGTAAA	3420
QY	3421	TTTTTTCAGCTATTTAGTGATATATTTTATCGGTGGGAATAAAAAATTTCTACTACAGAAAA	3480
Db	3421	TTTTTTCAGCTATTTAGTGATATATTTTATCGGTGGGAATAAAAAATTTCTACTACAGAAAA	3480
QY	3481	AAAAAAAAAAAAAAAAAAAAA 3501	
Db	3481	AAAAAAAAAAAAAAAAAAAAA 3501	
RESULT 4			
LOCUS	I40601	3501 bp	DNA linear PAT 13-MAY-1997
DEFINITION	Sequence 3 from patent US 5621090.		
ACCESSION	I40601		
VERSION	I40601.1 GI:2082893		
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 3501)		
AUTHORS	Lemischka, I.R.		
TITLE	Nucleic acids encoding soluble human FLK-2 extracellular domain		
JOURNAL	Patent: US 5621090-A 3 15-APR-1997;		
FEATURES	Location/Qualifiers		
source	1..3501		
BASE COUNT	1068 a	709 c	784 g 940 t
ORIGIN			
Query Match 100.0%; Score 3501; DB 6; Length 3501;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	CGAGGCGGCATCCGAGGCTGGGCGGCGCCCTGGGGGACCCCGGCTCCGAGGCCATG	60
Db	1	CGAGGCGGCATCCGAGGCTGGGCGGCGCCCTGGGGGACCCCGGCTCCGAGGCCATG	60
QY	61	CCGGCGTTGGCGCGGACCGGCGGCGCGCTGCTGCTGTTGTTTTTCTGCAATGATA	120
Db	61	CCGGCGTTGGCGCGGACCGGCGGCGCGCTGCTGCTGTTGTTTTTCTGCAATGATA	120
QY	121	TTTGGGACTATTACAATCAAGATCTGCTGTGATCAAGTGTGTTTTTAATCAATCAAG	180
Db	121	TTTGGGACTATTACAATCAAGATCTGCTGTGATCAAGTGTGTTTTTAATCAATCAAG	180

QY	181	AACAAATGATTCATCAGTGGGGAAGTTCATCATCATATCCATGGTATCAAGATCCCGGAA	240
Db	181	AACAAATGATTCATCAGTGGGGAAGTTCATCATCATATCCATGGTATCAAGATCCCGGAA	240
QY	241	GACCTCGGGTGTGGTGTGAGACCCAGAGCTCAGGACAGTGTACGAGAGCTCCCGCTGTG	300
Db	241	GACCTCGGGTGTGGTGTGAGACCCAGAGCTCAGGACAGTGTACGAGAGCTCCCGCTGTG	300
QY	301	GAAGTGGATGATCTGCTTCCATCACACTGCAAGTGTGTCGATGCCGCCAGGGAAACATT	360
Db	301	GAAGTGGATGATCTGCTTCCATCACACTGCAAGTGTGTCGATGCCGCCAGGGAAACATT	360
QY	361	TCCTGTCTCTGGGTCTTTTAAAGCAAGCTCCCTGAAATTCAGAGCCCAATTGTTGATTA	420
Db	361	TCCTGTCTCTGGGTCTTTTAAAGCAAGCTCCCTGAAATTCAGAGCCCAATTGTTGATTA	420
QY	421	AACAGAGGAGTGTGTTTCCATGGTCAATTTTGAATAATGACAGAAACCCAGCTGGAGAA	480
Db	421	AACAGAGGAGTGTGTTTCCA TGGTCAATTTTGAATAATGACAGAAACCCAGCTGGAGAA	480
QY	481	CTACTTTTATTCAGAGTGAAGCTACCAATTTACACAATATTTGTTTACAGTGAATATA	540
Db	481	CTACTTTTATTCAGAGTGAAGCTACCAATTTACACAATATTTGTTTACAGTGAATATA	540
QY	541	AATACCTGCTTTTACACATTAAGAAGACCTTACTTTAGAAAAATGGAACCCAGGACGC	600
Db	541	AATACCTGCTTTTACACATTAAGAAGACCTTACTTTAGAAAAATGGAACCCAGGACGC	600
QY	601	CTGGTCTGCATATCTGAGAGGCTTCCAGAGCGATCGTGAATGGTGTCTTTGCGATTCA	660
Db	601	CTGGTCTGCATATCTGAGAGGCTTCCAGAGCGATCGTGAATGGTGTCTTTGCGATTCA	660
QY	661	CAGGGGAAAGCTGTAAAGAAAGAAAGTCCAGCTGTGTTTAAAAAGGAGGAAAAAGTGT	720
Db	661	CAGGGGAAAGCTGTAAAGAAAGAAAGTCCAGCTGTGTTTAAAAAGGAGGAAAAAGTGT	720
QY	721	CATGAATTTATTTGGGACGACATGAAGTGTCTGCCAGAAATGAACTGGGCGAGGAAATGC	780
Db	721	CATGAATTTATTTGGGACGACATGAAGTGTCTGCCAGAAATGAACTGGGCGAGGAAATGC	780
QY	781	ACAGGCTGTTCACATAGATCTAAATCAAACTCCTCAGACCAATTTGCAATTAATTT	840
Db	781	ACCAGGCTGTTCACATAGATCTAAATCAAACTCCTCAGACCAATTTGCAATTAATTT	840
QY	841	CTTAAAGTGGGAAACCCCTTATGATGAAGTGCAGAGCTGTTTCAATGGAACCATGATTC	900
Db	841	CTTAAAGTGGGAAACCCCTTATGATGAAGTGCAGAGCTGTTTCAATGGAACCATGATTC	900
QY	901	GGGCTCACCTGGGAAATTTAGAAAAACAAGCACTCGAGGAGGCAACTACTTTGAGATGAT	960
Db	901	GGGCTCACCTGGGAAATTTAGAAAAACAAGCACTCGAGGAGGCAACTACTTTGAGATGAT	960
QY	961	ACCTATTCAACAAACAGAACTATGATACGGAATCTGTTTCTTTTGTATCATCATGCGCA	1020
Db	961	ACCTATTCAACAAACAGAACTATGATACGGAATCTGTTTCTTTTGTATCATCATGCGCA	1020
QY	1021	AGAAACGACACCGGATCTACACTTGTCTTCAAGCATCCAGTCAATCATGCTTTG	1080
Db	1021	AGAAACGACACCGGATCTACACTTGTCTTCAAGCATCCAGTCAATCATGCTTTG	1080
QY	1081	GTTTACCATCGTAGGAAGGATTTATAAATGCTACCAATTTCAAGTGAAGATTATGAAT	1140
Db	1081	GTTTACCATCGTAGGAAGGATTTATAAATGCTACCAATTTCAAGTGAAGATTATGAAT	1140
QY	1141	GACCAATATGAAGATTTTGTCTTCTGTCAGGTTTAAAGCCTACCAAAATCAGATGT	1200
Db	1141	GACCAATATGAAGATTTTGTCTTCTGTCAGGTTTAAAGCCTACCAAAATCAGATGT	1200
QY	1201	ACGTGACCTTCTCTGAAAAATCAATTTCTTGTGAGCAAAAGGGTCTTGTATAACGATAC	1260
Db	1201	ACGTGACCTTCTCTGAAAAATCAATTTCTTGTGAGCAAAAGGGTCTTGTATAACGATAC	1260

1261 AGCATATCCAAAGTTTGGCAATCATAGCACACAGCCAGGAGAAATATATATATTCATTCAGAGAA 1320  
1261 AGCATATCCAAAGTTTGGCAATCATAGCACACAGCCAGGAGAAATATATATATTCATTCAGAGAA 1320  
1321 AATGATGATGCCCAATTTTACCAAAATGTTTCACCTGAAATATAGAGAGAAACCTCAAGTG 1380  
1321 AATGATGATGCCCAATTTTACCAAAATGTTTCACCTGAAATATAGAGAGAAACCTCAAGTG 1380  
1381 CTCGCAAGAACATCGGCAAGTCAGGGGCTCCTGTTTCTCGGATGGATACCCATACCATCT 1440  
1381 CTCGCAAGAACATCGGCAAGTCAGGGGCTCCTGTTTCTCGGATGGATACCCATACCATCT 1440  
1441 TGGACCTGGAAGAGTGTTCAGACAAGTCTCCCAAATGTCGACAGAGATCAAGAGGA 1500  
1441 TGGACCTGGAAGAGTGTTCAGACAAGTCTCCCAAATGTCGACAGAGATCAAGAGGA 1500  
1501 GTCTGGAATAGAAAGGCTTAACAGAAAGTGTTCGACAGTGGGTGTCGACGACTACTCTA 1560  
1501 GTCTGGAATAGAAAGGCTTAACAGAAAGTGTTCGACAGTGGGTGTCGACGACTACTCTA 1560  
1561 AACATGAGTGAAGCCATAAAAGGTTCTCTGCTCAAGTGTGTCGATACAAATTCCTTTGGC 1620  
1561 AACATGAGTGAAGCCATAAAAGGTTCTCTGCTCAAGTGTGTCGATACAAATTCCTTTGGC 1620  
1621 ACATCTGTGAGACGATCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCACAAGACAAC 1680  
1621 ACATCTGTGAGACGATCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCACAAGACAAC 1680  
1681 ATCTCAATTCATGCAACAAATGTTGTTGTTCTCTCTCAATGTCGTTTAAACCTTGCTA 1740  
1681 ATCTCAATTCATGCAACAAATGTTGTTGTTCTCTCTCAATGTCGTTTAAACCTTGCTA 1740  
1741 ATTGTGTCAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
1741 ATTGTGTCAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
1801 ACCGGTCTCTCAGATAATAGTACTCTACGTTGATTTTCAGAGAAATATGAATATGATCTC 1860  
1801 ACCGGTCTCTCAGATAATAGTACTCTACGTTGATTTTCAGAGAAATATGAATATGATCTC 1860  
1861 AAATGGAGTTTCCAGAGAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920  
1861 AAATGGAGTTTCCAGAGAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920  
1921 GGAAGAGTATGAACGCAACAGCTTATGGAATTTAGCAAAAACAGAGTCTCAATCCAGGTT 1980  
1921 GGAAGAGTATGAACGCAACAGCTTATGGAATTTAGCAAAAACAGAGTCTCAATCCAGGTT 1980  
1981 GCCGTCAAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCAAGAA 2040  
1981 GCCGTCAAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCAAGAA 2040  
2041 CTCAGATGATGACCCAGCTGGGAGCCACGAGNATATTTGAACTCTGGGGGGCGTGC 2100  
2041 CTCAGATGATGACCCAGCTGGGAGCCACGAGNATATTTGAACTCTGGGGGGCGTGC 2100  
2101 ACACTGTCAGGACCAATTTACTGATTTTGAATFACCTGTTGCTATGTTGATCTTCTCAAC 2160  
2101 ACACTGTCAGGACCAATTTACTGATTTTGAATFACCTGTTGCTATGTTGATCTTCTCAAC 2160  
2161 TATCTAAGAGTAAAGAGAAAATTTTCAGGACTTGGACAGAGATTTTCAAGGAAACAC 2220  
2161 TATCTAAGAGTAAAGAGAAAATTTTCAGGACTTGGACAGAGATTTTCAAGGAAACAC 2220  
2221 AATTTCAAGTTTTCACCCATCTTCCAAATTCATCCTCAATTCAGATGCTGGTTCAAGA 2280  
2221 AATTTCAAGTTTTCACCCATCTTCCAAATTCATCCTCAATTCAGATGCTGGTTCAAGA 2280  
2281 GAAGTTTCAGATACACCCGGCTCGGATCAATCTCAGGGCTTCATGGGAATTCATTTTCA 2340  
2281 GAAGTTTCAGATACACCCGGCTCGGATCAATCTCAGGGCTTCATGGGAATTCATTTTCA 2340  
2341 TCTGAAGATGAAATTTGAATATGAAAAACCAAAAAAGGCTGGGAAAGAGAGGAGGACTTGAAT 2400

2341 TCTGAAGATGAAATTTGAATATGAAAAACCAAAAAAGGCTGGAGAAAGAGGAGGACTTGAAT 2400  
2401 GTGCTTACATTTGAAGATCTTCTTTGCTTTGCAATATCAAGTTTCCCAAGGAATGGAATTT 2460  
2401 GTGCTTACATTTGAAGATCTTCTTTGCTTTGCAATATCAAGTTTCCCAAGGAATGGAATTT 2460  
2461 CTGGAATTTAAGTCGTGTGTTTCAGAGACCTGGCCGCGAGAAACGTGCTTGTCAACCAC 2520  
2461 CTGGAATTTAAGTCGTGTGTTTCAGAGACCTGGCCGCGAGAAACGTGCTTGTCAACCAC 2520  
2521 GGGAAAGTCGTGAAGATATGTCATTTGGATTGGCTCGAGATATCATGAGTGAATTCACAAC 2580  
2521 GGGAAAGTCGTGAAGATATGTCATTTGGATTGGCTCGAGATATCATGAGTGAATTCACAAC 2580  
2581 TATGTTGTGAGGGGCAATGCCCTCTGCTGTAATAATGATGGCTGGCCCGGAAAAGCCTGTTT 2640  
2581 TATGTTGTGAGGGGCAATGCCCTCTGCTGTAATAATGATGGCTGGCCCGGAAAAGCCTGTTT 2640  
2641 GAAGGCATCTACACCAATTAAGAGTGTCTGCTCATATGGAATATTTACTGTGGGAAATC 2700  
2641 GAAGGCATCTACACCAATTAAGAGTGTCTGCTCATATGGAATATTTACTGTGGGAAATC 2700  
2701 TTCTCACTTGGTGTGAATCCTTACCTGGCATTTCCGGTTGATCGTAACTTCTACAAACTG 2760  
2701 TTCTCACTTGGTGTGAATCCTTACCTGGCATTTCCGGTTGATCGTAACTTCTACAAACTG 2760  
2761 ATTCAAAATGGATTTAAAATGATCAGCCATTTTATGCTACAGAAAGAAATATACATTATA 2820  
2761 ATTCAAAATGGATTTAAAATGATCAGCCATTTTATGCTACAGAAAGAAATATACATTATA 2820  
2821 ATGCAATCTCTGCTGGGCTTTTGACTCAAGGAAAACGGCCATCTCTTCCCTAAATTTGACTTCG 2880  
2821 ATGCAATCTCTGCTGGGCTTTTGACTCAAGGAAAACGGCCATCTCTTCCCTAAATTTGACTTCG 2880  
2881 TTTTTHAGATGTCAGCTGGCAGATGTCAGAAAGCCGATGATCAGATGTTGGATGGCGGT 2940  
2881 TTTTTHAGATGTCAGCTGGCAGATGTCAGAAAGCCGATGATCAGATGTTGGATGGCGGT 2940  
2941 GTTTCGGATGTCCTCAACCTTACCAAAAACAGGCGACCTTTTACAGAGAGATGGAATTTG 3000  
2941 GTTTCGGATGTCCTCAACCTTACCAAAAACAGGCGACCTTTTACAGAGAGATGGAATTTG 3000  
3001 GGGCTACTCTCTCCGAGGCTCAGGTGCGAAGATTCGTAGAGGAAACAAATTTAGTTTAAAG 3060  
3001 GGGCTACTCTCTCCGAGGCTCAGGTGCGAAGATTCGTAGAGGAAACAAATTTAGTTTAAAG 3060  
3061 ACTTCATCCCTCAACCTATCCCTAAACAGGCTGTAGATTACCAAAAACAGATTTAATTTTCT 3120  
3061 ACTTCATCCCTCAACCTATCCCTAAACAGGCTGTAGATTACCAAAAACAGATTTAATTTTCT 3120  
3121 CACTAAAAGAAAATCTATTATCAACTGCTCTTCCAGAGACTTTTCTTAGAAGCCGTCT 3180  
3121 CACTAAAAGAAAATCTATTATCAACTGCTCTTCCAGAGACTTTTCTTAGAAGCCGTCT 3180  
3181 GCCTTTACTCTGTTTTCAAAAGGGAATTTTGTAAAAATCAAAATCATCTGTCACAAGGCAG 3240  
3181 GCCTTTACTCTGTTTTCAAAAGGGAATTTTGTAAAAATCAAAATCATCTGTCACAAGGCAG 3240  
3241 GAGGAGCTGATATGAACCTTTATTTGGAGCATTTGATCTGCAATCCAAAGGCCCTTCTCAGGCCG 3300  
3241 GAGGAGCTGATATGAACCTTTATTTGGAGCATTTGATCTGCAATCCAAAGGCCCTTCTCAGGCCG 3300  
3301 GCTTGAGTGAATTTGTGTACCTGAGTACAGTATATCTTCTGTAATACATAAAAACAAAAGC 3360  
3301 GCTTGAGTGAATTTGTGTACCTGAGTACAGTATATCTTCTGTAATACATAAAAACAAAAGC 3360  
3361 ATTTTGTAGAGAGAGCTAATATGATTTTAAAGTCTATGTTTAAAAATAATATGTAAA 3420  
3361 ATTTTGTAGAGAGAGCTAATATGATTTTAAAGTCTATGTTTAAAAATAATATGTAAA 3420  
3421 TTTTTCAGCTATTTAGTGATATATTTTATGGGTGGGAATAAAAATTTCTTCTACAGAAAAA 3480

Db	3421	TTTTCTAGCTATTTAGTGTATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAAA	3480
Qy	3481	AAAAAAAAAAAAAAAAAAAAA 3501	
Db	3481	AAAAAAAAAAAAAAAAAAAAA 3501	
RESULT 5			
LOCUS	AX695522	3475 bp	DNA linear PAT 31-MAR-2003
DEFINITION	Sequence 1149 from Patent WO03008583.		
ACCESSION	AX695522		
VERSION	AX695522.1	GI:29418674	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Morris,D.W. and Engelhard,E.K.		
TITLE	Novel compositions and methods for cancer		
JOURNAL	Patent: WO 03008583-A 1149 30-JAN-2003;		
FEATURES	Sagres Discovery (US)		
source	Location/Qualifiers		
	1..3475		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
BASE COUNT	1042 a	709 c	784 g 940 t
ORIGIN			
Query Match	99.3%; Score 3475; DB 6; Length 3475;		
Best Local Similarity	100.0%; Pred. No. 0;		
Matches 3475;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1	CGAGGCGGCATCCGAGGGCTGGGCGGCGCCCTCGGGGACCCCGGGCTCCGGAGGCCATG	60
Db	1	CGAGGCGGCATCCGAGGGCTGGGCGGCGCCCTCGGGGACCCCGGGCTCCGGAGGCCATG	60
Qy	61	CCGGCTGTGGCGCGGACCGGGCACCGTGCCTGCTCGTGTGTTTCTTCTGCAATGATA	120
Db	61	CCGGCTGTGGCGCGGACCGGGCACCGTGCCTGCTCGTGTGTTTCTTCTGCAATGATA	120
Qy	121	TTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTAAATCAATCAATAAG	180
Db	121	TTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTAAATCAATCAATAAG	180
Qy	181	AACAATGATTCAATCAGTGGGAATCATATCATATCCATGGTATCAGAAATCCCGGAA	240
Db	181	AACAATGATTCAATCAGTGGGAATCATATCATATCCATGGTATCAGAAATCCCGGAA	240
Qy	241	GACCTGGGTGTGGTTGAGACCCAGAGCTCAGGGACAGTGTACGAGCTGCCGCTGTG	300
Db	241	GACCTGGGTGTGGTTGAGACCCAGAGCTCAGGGACAGTGTACGAGCTGCCGCTGTG	300
Qy	301	GAAGTGATGTATCTGCTTCCATCACACTGCAAGTGTGTCGATGCCAGGGAAACATT	360
Db	301	GAAGTGATGTATCTGCTTCCATCACACTGCAAGTGTGTCGATGCCAGGGAAACATT	360
Qy	361	TCCTGTCTCTGGGTCTTTAAGCACAGCTTCCGTGAATTCGACGCCACATTTTGAATTACA	420
Db	361	TCCTGTCTCTGGGTCTTTAAGCACAGCTTCCGTGAATTCGACGCCACATTTTGAATTACA	420
Qy	421	AACAGAGGAGTGTGTTCCATGGTCAATTTGAAATGACAGAAACCCAGCTGGAGATAC	480
Db	421	AACAGAGGAGTGTGTTCCATGGTCAATTTGAAATGACAGAAACCCAGCTGGAGATAC	480
Qy	481	CTACTTTTATTCAGAGTGAAGCTACCAATTACAAATATTTGTTTACAGTGAGTATAAGA	540
Db	481	CTACTTTTATTCAGAGTGAAGCTACCAATTACAAATATTTGTTTACAGTGAGTATAAGA	540
Qy	541	AATACCCCTGCTTTTACATTAAGGAAGACCTTACTTTTAAAAATGGAAAAACCGAGACGCC	600

```
QY 1681 ATCTCAATTCATGCAACAAATGGGTGTTGTCTCTCTCTCAATGTCGTTTAAACCCCTGCTA 1740
DB |||||||
DB 1681 ATCTCAATTCATGCAACAAATGGGTGTTGTCTCTCTCTCAATGTCGTTTAAACCCCTGCTA 1740
QY 1741 ATTTGTCACAAGTACAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800
DB |||||||
DB 1741 ATTTGTCACAAGTACAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800
QY 1801 ACCGGCTCCTCAGATAATAGTACTCTTACCTGTAATTTACAGAGAAATATGAATATGATCTC 1860
DB |||||||
DB 1801 ACCGGCTCCTCAGATAATAGTACTCTTACCTGTAATTTACAGAGAAATATGAATATGATCTC 1860
QY 1861 AATGGGAGTTTCCAGAGAGAAATTTAGAGTTTGGGAAAGGTACTAGAGATCAGGTGCTTTT 1920
DB |||||||
DB 1861 AATGGGAGTTTCCAGAGAGAAATTTAGAGTTTGGGAAAGGTACTAGAGATCAGGTGCTTTT 1920
QY 1921 GGAAGAGTGTATGACGCAACAGCTTATGGAATTAAGCAAGCAAGAGAGTCTCAATCCAGGTT 1980
DB |||||||
DB 1921 GGAAGAGTGTATGACGCAACAGCTTATGGAATTAAGCAAGCAAGAGAGTCTCAATCCAGGTT 1980
QY 1981 GCCGTCAAAATGCTGAAAGAAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGCAAA 2040
DB |||||||
DB 1981 GCCGTCAAAATGCTGAAAGAAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGCAAA 2040
QY 2041 CTCAAGATGATGACCCAGCTGGGAAAGCAACGAGAAATATGTGAACTGTGGGGCGGTGC 2100
DB |||||||
DB 2041 CTCAAGATGATGACCCAGCTGGGAAAGCAACGAGAAATATGTGAACTGTGGGGCGGTGC 2100
QY 2101 ACATGTCAGGACCAATTTACTGATTTTGAATTAATCTGTTGCTATGTTGATCTCTCAAC 2160
DB |||||||
DB 2101 ACATGTCAGGACCAATTTACTGATTTTGAATTTGAACTGTTGCTATGTTGATCTCTCAAC 2160
QY 2161 TATCTAAGAAAGTAAAGAGAAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAAACAC 2220
DB |||||||
DB 2161 TATCTAAGAAAGTAAAGAGAAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAAACAC 2220
QY 2221 AATTTCAAGTTTATCCCACTTTTCCAAATCAATCCAAATCCAGATGCTGCTGGTTCGAAGA 2280
DB |||||||
DB 2221 AATTTCAAGTTTATCCCACTTTTCCAAATCAATCCAAATCCAGATGCTGCTGGTTCGAAGA 2280
QY 2281 GAAGTTCAAGTACACCCGGACTCCGATCAATCTCAGGGCTTCATGGGAATTCATTTCCAC 2340
DB |||||||
DB 2281 GAAGTTCAAGTACACCCGGACTCCGATCAATCTCAGGGCTTCATGGGAATTCATTTCCAC 2340
QY 2341 TCTGAAGATGAAATTTGAATATGAAAAACCAAAAGGCTGGAAAGAGAGAGAGACTTGAAT 2400
DB |||||||
DB 2341 TCTGAAGATGAAATTTGAATATGAAAAACCAAAAGGCTGGAAAGAGAGAGAGACTTGAAT 2400
QY 2401 GTGCTTACATTTGAAGATCTTTCTTTGCTTTCATATCAAGTTGGCCAAAGGAATGGAATTT 2460
DB |||||||
DB 2401 GTGCTTACATTTGAAGATCTTTCTTTGCTTTCATATCAAGTTGGCCAAAGGAATGGAATTT 2460
QY 2461 CTGGAATTTAAGTCGTGTGTTTACAGAGACCTGCGCCGACGAGAGAGTCTGTCAACCCAC 2520
DB |||||||
DB 2461 CTGGAATTTAAGTCGTGTGTTTACAGAGACCTGCGCCGACGAGAGAGTCTGTCAACCCAC 2520
QY 2521 GGGAAAGTGGTGAAGATATGTGACTTTTGGATTTGGCTCGAGATATCATGAGTATTCCCAAC 2580
DB |||||||
DB 2521 GGGAAAGTGGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTATTCCCAAC 2580
QY 2581 TATGTTGTGAGGGGCAATGCCCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640
DB |||||||
DB 2581 TATGTTGTGAGGGGCAATGCCCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640
QY 2641 GAAGGCATCTACACCAATTAAGAGTGTATGCTGGTTCATATGGAATATTAATCTGCGGAATC 2700
DB |||||||
DB 2641 GAAGGCATCTACACCAATTAAGAGTGTATGCTGGTTCATATGGAATATTAATCTGCGGAATC 2700
QY 2701 TTCTCCTGTTGTGTAATCCCTTACCCCTGGCAATCCCGTTGATGCTTAACCTTACAAAACCTG 2760
DB |||||||
DB 2701 TTCTCCTGTTGTGTAATCCCTTACCCCTGGCAATCCCGTTGATGCTTAACCTTACAAAACCTG 2760
```

```
QY 2761 ATTCAAAATCGATTAAATGGATCCAGCATTTTATGCTACAGAGAAATATACATTATA 2820
DB |||||||
DB 2761 ATTCAAAATCGATTAAATGGATCCAGCATTTTATGCTACAGAGAAATATACATTATA 2820
QY 2821 ATGCAATCCCTGCTGGGCTTTTGAATCAAGGAAACGGCCATCTTCCCTAAATTTGACTTCG 2880
DB |||||||
DB 2821 ATGCAATCCCTGCTGGGCTTTTGAATCAAGGAAACGGCCATCTTCCCTAAATTTGACTTCG 2880
QY 2881 TTTTATAGATGTGAGCTGGCAGATGCAAGAAAGCGATGTATCAGAAATGTGGATGGCCGT 2940
DB |||||||
DB 2881 TTTTATAGATGTGAGCTGGCAGATGCAAGAAAGCGATGTATCAGAAATGTGGATGGCCGT 2940
QY 2941 GTTTTCGGAATGTCCTCAACCTTCAACCAAAACAGGGCGA CTTTTCAGCAGAGATGGATTTG 3000
DB |||||||
DB 2941 GTTTTCGGAATGTCCTCAACCTTCAACCAAAACAGGGCGA CTTTTCAGCAGAGATGGATTTG 3000
QY 3001 GGGCTACTCTCTCCGAGGCTCAGGTCGAAAGATTCGTAGAGGAAACAATTTAGTTTAAAGG 3060
DB |||||||
DB 3001 GGGCTACTCTCTCCGAGGCTCAGGTCGAAAGATTCGTAGAGGAAACAATTTAGTTTAAAGG 3060
QY 3061 ACTTCATCCCTCCACCTTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTAAATTCAT 3120
DB |||||||
DB 3061 ACTTCATCCCTCCACCTTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTAAATTCAT 3120
QY 3121 CACTAAAGAAAAATCTATTATCAACTGCTGCTTCCAGAGACTTTTCTCTAGAACCCGTCT 3180
DB |||||||
DB 3121 CACTAAAGAAAAATCTATTATCAACTGCTGCTTCCAGAGACTTTTCTCTAGAACCCGTCT 3180
QY 3181 GCGTTTACTCTGTGTTTCAAAGGAGCTTTTGTAAAAATCAAAATCATCTGTCACAAGGCAG 3240
DB |||||||
DB 3181 GCGTTTACTCTGTGTTTCAAAGGAGCTTTTGTAAAAATCAAAATCATCTGTCACAAGGCAG 3240
QY 3241 GAGGAGCTGATAATGAACCTTTATTGGAGCATTTGATCTGCATCCAAAGGCCCTTCTCAGGCCG 3300
DB |||||||
DB 3241 GAGGAGCTGATAATGAACCTTTATTGGAGCATTTGATCTGCATCCAAAGGCCCTTCTCAGGCCG 3300
QY 3301 GCTTGAGTGAATGTGTACCTGAAGTACAGTATATCTTGTGAAATACATAAAACAAGC 3360
DB |||||||
DB 3301 GCTTGAGTGAATGTGTACCTGAAGTACAGTATATCTTGTGAAATACATAAAACAAGC 3360
QY 3361 ATTTTGTCTAAGGAGAACTAATGATTTTAAAGTCTATGTTTAAATTAATATGTAATA 3420
DB |||||||
DB 3361 ATTTTGTCTAAGGAGAACTAATGATTTTAAAGTCTATGTTTAAATTAATATGTAATA 3420
QY 3421 TTTTTCAGCTATTATGATGATATATTTATGGTGGGAATAAAATTTCTACTACAG 3475
DB |||||||
DB 3421 TTTTTCAGCTATTATGATGATATATTTATGGTGGGAATAAAATTTCTACTACAG 3475
```

## RESULT 6

```
144733 LOCUS 144733 3475 bp DNA linear PAT 07-OCT-1997
DEFINITION Sequence 3 from patent US 5635388.
ACCESSION 144733
VERSION 144733.1 GI:2469446
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified
REFERENCE 1 (bases 1 to 3475)
AUTHORS Bennett,B.D., Broz,S.D., Matthews,W. and Zeigler,F.C.
TITLE Agonist antibodies against the flk2/flt3 receptor and uses thereof
JOURNAL Patent: US 5635388-A 3 03-JUN-1997;
FEATURES Location/Qualifiers
source 1..3475
BASE COUNT 1042 a 709 c 784 g 940 t
ORIGIN
```

```
Query Match 99.3%; Score 3475; DB 6; Length 3475;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	1	CGAGCGGCATCCGAGGCTGGCGCGCCCTGGGGGACCCCGGGCTCCGGAGGCCATG	60	Db	1081	GTATACCATCGTAGGAAGGGATTTATAATGCTACCAATTCAGTGAAGATTATGAAT	1140
Db	1	CGAGCGGCATCCGAGGCTGGCGCGCCCTGGGGGACCCCGGGCTCCGGAGGCCATG	60	Qy	1141	GACCAATATGAAGAGTTTGTCTGTCTAGGTTTAAAGCTACCCACAAATCAGATGT	1200
Qy	61	CCGCGTTGGCGCGACGCGGGGACCCGCTGCGTCTGTGTTTCTGCAATGATA	120	Db	1141	GACCAATATGAAGAGTTTGTCTGTCTAGGTTTAAAGCTACCCACAAATCAGATGT	1200
Db	61	CCGCGTTGGCGCGACGCGGGGACCCGCTGCGTCTGTGTTTCTGCAATGATA	120	Qy	1201	ACGTGGACCTTCTCTCGAANAATCAATTCCTTGTGAGCAAAAGGGTCTTGATTAACGGATAC	1260
Qy	121	TTTGGGACTATTACAAATCAAGATCTGCTGTGTGATCAAGTGTGTTTAATCAATCAAG	180	Db	1201	ACGTGGACCTTCTCTCGAANAATCAATTCCTTGTGAGCAAAAGGGTCTTGATTAACGGATAC	1260
Db	121	TTTGGGACTATTACAAATCAAGATCTGCTGTGTGATCAAGTGTGTTTAATCAATCAAG	180	Qy	1261	AGCATATCCAAGTTTGTGCAATCAAGCACACGAGCAGGAGAAATATATATTCATTCAGAA	1320
Qy	181	AACAAATGATTCATCAGTGGGGAAGTCAATCATATCCATGGGTATCAGAAATCCCGGAA	240	Db	1261	AGCATATCCAAGTTTGTGCAATCAAGCACACGAGCAGGAGAAATATATATTCATTCAGAA	1320
Db	181	AACAAATGATTCATCAGTGGGGAAGTCAATCATATCCATGGGTATCAGAAATCCCGGAA	240	Qy	1321	AATGATGATGCCCAATTTTACCAAAATGTTTCCGCTGAAATATAAGAAAGGAAACCTCAAGTG	1380
Qy	241	GACCTCGGCTGTGCTTGAGACCCGAGAGCTCAGGACAGTGTACCAAGCTGCCGCTGTG	300	Db	1321	AATGATGATGCCCAATTTTACCAAAATGTTTCCGCTGAAATATAAGAAAGGAAACCTCAAGTG	1380
Db	241	GACCTCGGCTGTGCTTGAGACCCGAGAGCTCAGGACAGTGTACCAAGCTGCCGCTGTG	300	Qy	1381	CTCGAGAAAGCATCGGCAAGTCAAGCGTCTGTGTTTCTCGATGGATACCCATTAATCT	1440
Qy	301	GAAATGATGATCTGCTTCCATCACTGCAAGTGTGCTGATGCTCCGAGGGAACATT	360	Db	1381	CTCGAGAAAGCATCGGCAAGTCAAGCGTCTGTGTTTCTCGATGGATACCCATTAATCT	1440
Db	301	GAAATGATGATCTGCTTCCATCACTGCAAGTGTGCTGATGCTCCGAGGGAACATT	360	Qy	1441	TGGACCTGGAAAGAGTGTTCAGACAAGTCTCCCACTGCACAGAAAGATCAAGAAGGA	1500
Qy	361	TCCTGTCTCGGCTTTTAAAGCAGAGTCCCTGAAATTCGACAGCCACATTTTGAATTA	420	Db	1441	TGGACCTGGAAAGAGTGTTCAGACAAGTCTCCCACTGCACAGAAAGATCAAGAAGGA	1500
Db	361	TCCTGTCTCGGCTTTTAAAGCAGAGTCCCTGAAATTCGACAGCCACATTTTGAATTA	420	Qy	1501	GTCTCGAATAGAAAGGCTTAAACAGAAAGTGTGAGCAAGTGGTTCGAGCAGTACTCTA	1560
Qy	421	AACAGAGAGTGTGTTTCCATGCTTAAATGAAATGACAGAAACCCAGCTGGGAATAC	480	Db	1501	GTCTCGAATAGAAAGGCTTAAACAGAAAGTGTGAGCAAGTGGTTCGAGCAGTACTCTA	1560
Db	421	AACAGAGAGTGTGTTTCCATGCTTAAATGAAATGACAGAAACCCAGCTGGGAATAC	480	Qy	1561	AACATGATGAAGCATATAAAGGTTCTGTGTCAGTCTGTGCATACAATTCCTTGGC	1620
Qy	481	CTACTTTTATTCAGAGTGAAGTACCAATTAACAAATTTTGTGTTTACAGTGAATAGA	540	Db	1561	AACATGATGAAGCATATAAAGGTTCTGTGTCAGTCTGTGCATACAATTCCTTGGC	1620
Db	481	CTACTTTTATTCAGAGTGAAGTACCAATTAACAAATTTTGTGTTTACAGTGAATAGA	540	Qy	1621	ACATCTTGTGAGACGATCTTTTAAACTCTCCAGGCCCTTCCCTTTCATCAAGACAC	1680
Qy	541	AATACCTGCTTTACATTAAGAGACCTTACTTTTAAAGAAATGGAATAACAGAGGCC	600	Db	1621	ACATCTTGTGAGACGATCTTTTAAACTCTCCAGGCCCTTCCCTTTCATCAAGACAC	1680
Db	541	AATACCTGCTTTACATTAAGAGACCTTACTTTTAAAGAAATGGAATAACAGAGGCC	600	Qy	1681	ATCTCATTTCTATGCAACAAATTTGGTGTGTTCTCTCTTCTTCTTCTTCTTCTTCTA	1740
Qy	601	CTGCTGCTATCTGAGAGGTTTCCAGAGCGATCGTGGAAATGGTCTTTGGATTC	660	Db	1681	ATCTCATTTCTATGCAACAAATTTGGTGTGTTCTCTCTTCTTCTTCTTCTTCTTCTA	1740
Db	601	CTGCTGCTATCTGAGAGGTTTCCAGAGCGATCGTGGAAATGGTCTTTGGATTC	660	Qy	1741	ATTTGTCAAGTCAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGGTG	1800
Qy	661	CAGGGGAAAGCTGTAAAGAAAGTCCAGCTGTTTAAAGAGGAGGAAAGTGTCT	720	Db	1741	ATTTGTCAAGTCAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGGTG	1800
Db	661	CAGGGGAAAGCTGTAAAGAAAGTCCAGCTGTTTAAAGAGGAGGAAAGTGTCT	720	Qy	1801	ACCGCTCTCTCAGATAATGAGTACTTCTACGTTGATTTTACAGAAATATGAATATCTC	1860
Qy	721	CATGAATTTATTTGGGACGACATAAGGTGCTGCGCAAGAAATGAATGGGAGGAAATGC	780	Db	1801	ACCGCTCTCTCAGATAATGAGTACTTCTACGTTGATTTTACAGAAATATGAATATCTC	1860
Db	721	CATGAATTTATTTGGGACGACATAAGGTGCTGCGCAAGAAATGAATGGGAGGAAATGC	780	Qy	1861	AAATGGGAGTTTCCAGAGAAATTTAGAGTTTGGAAAGGTACTAGGATCAGGTCTTTT	1920
Qy	781	ACGAGGCTGTTCAATATAGATCTAAATCAAACTCTCAGACCAATTCGCCAATATTT	840	Db	1861	AAATGGGAGTTTCCAGAGAAATTTAGAGTTTGGAAAGGTACTAGGATCAGGTCTTTT	1920
Db	781	ACGAGGCTGTTCAATATAGATCTAAATCAAACTCTCAGACCAATTCGCCAATATTT	840	Qy	1921	GGAAAAGTGAAGCAAGCAAGCTTTATGGAATTAGCAAAACAGAGGTCTCAATCCAGGT	1980
Qy	841	CTTAAAGTGGGGAACCCCTTATGATAGGTGCAAGCTGTTTCAATCGATGATTC	900	Db	1921	GGAAAAGTGAAGCAAGCAAGCTTTATGGAATTAGCAAAACAGAGGTCTCAATCCAGGT	1980
Db	841	CTTAAAGTGGGGAACCCCTTATGATAGGTGCAAGCTGTTTCAATCGATGATTC	900	Qy	1981	GCCTCAAAATGCTGAAAGAAAAAGCAGACAGCTCTGAAAAGAGGCACTCATGTGAGAA	2040
Qy	901	GGGCTACCTGGGAATTAGAAAAACAAAGCACTCGAGGAGGCAACTCTTTGAGATGAT	960	Db	1981	GCCTCAAAATGCTGAAAGAAAAAGCAGACAGCTCTGAAAAGAGGCACTCATGTGAGAA	2040
Db	901	GGGCTACCTGGGAATTAGAAAAACAAAGCACTCGAGGAGGCAACTCTTTGAGATGAT	960	Qy	2041	CTCAGATGATGCCAGCTGGGAGCCACAGAAATATTGAACTCTCTGGGGGGTGC	2100
Qy	961	ACCTATTCAACAAACAGAACTATGATACGATCTGTTTGTGTTTGTATCATCATGTC	1020	Db	2041	CTCAGATGATGCCAGCTGGGAGCCACAGAAATATTGAACTCTCTGGGGGGTGC	2100
Db	961	ACCTATTCAACAAACAGAACTATGATACGATCTGTTTGTGTTTGTATCATCATGTC	1020	Qy	2101	ACACTGTGAGGACCAATTTTACTGATTTTGAATTAATCTGTTGTTGTTGTTGTTGTTG	2160
Qy	1021	AGAAACGACACCGGATCTACACTGTTTCTTCAAGAGCATCCAGTCAATCAGCTTTG	1080	Db	2101	ACACTGTGAGGACCAATTTTACTGATTTTGAATTAATCTGTTGTTGTTGTTGTTGTTG	2160
Db	1021	AGAAACGACACCGGATCTACACTGTTTCTTCAAGAGCATCCAGTCAATCAGCTTTG	1080	Qy	2161	TATCTAAGAGTAAAGAGAAAAATTTTCAAGGACTTTGACAGAGATTTTCAAGGAACAC	2220
Qy	1081	GTTTACCATCGTAGGAAGGGATTTATAATGCTACCAATTCAGTGAAGATTATGAAT	1140				



2161	Db	TATCTAAGAGTAAAGAGAAAAATTTTCA CAGGACTTGGACAGAGATTTTTCAGGAACAC	2221	Qy	AAATTTACAGTTTTTACCCCACTTTTCCAATTCATCAATCAAAATTCAGCATGCTGTGTTCAAGA
2221	Db	AAATTTACAGTTTTTACCCCACTTTTCCAATTCATCAATCAAAATTCAGCATGCTGTGTTCAAGA	2281	Qy	GAAGTTTCAGATACACCCGGACTCGGATCAAAATCTCAGGGCTTTCATGGGAATTCATTTTCAC
2281	Db	GAAGTTTCAGATACACCCGGACTCGGATCAAAATCTCAGGGCTTTCATGGGAATTCATTTTCAC	2341	Qy	TCTGAAGATGAAAATTGAATATGAAAAACAAAAAGGCTGGAAGAAAGGAGGAGGACTTCTGAAT
2341	Db	TCTGAAGATGAAAATTGAATATGAAAAACAAAAAGGCTGGAAGAAAGGAGGAGGACTTCTGAAT	2401	Qy	GTGCTTACATTTGGAAGATCTTTCTTTGTCTTTCATATCAAGTTGTCAAAGGAAATGGAATTT
2401	Db	GTGCTTACATTTGGAAGATCTTTCTTTGTCTTTCATATCAAGTTGTCAAAGGAAATGGAATTT	2461	Qy	CTGGAAATTTAAGTCTGTGTGTTTCAGAGACCTGGCCGCGCAGGAACTGCTTTGTCAACCAC
2461	Db	CTGGAAATTTAAGTCTGTGTGTTTCAGAGACCTGGCCGCGCAGGAACTGCTTTGTCAACCAC	2521	Qy	GGGAAAGTGTGTAAGATATGTGACTTTTGGATTTGCTCGAGATATCATGAGTGAATTCACAC
2521	Db	GGGAAAGTGTGTAAGATATGTGACTTTTGGATTTGCTCGAGATATCATGAGTGAATTCACAC	2581	Qy	TATGTTCTCAGGGGCAATGCCGCTCTGCTGTAAATGCGATGGCCCGCAGAAAGCCTGTTT
2581	Db	TATGTTCTCAGGGGCAATGCCGCTCTGCTGTAAATGCGATGGCCCGCAGAAAGCCTGTTT	2641	Qy	GAAGGCATCTACACCAATTAAGAGTGATGCTCTGGTCTATATGGAATATTTACTGTGGGAAATC
2641	Db	GAAGGCATCTACACCAATTAAGAGTGATGCTCTGGTCTATATGGAATATTTACTGTGGGAAATC	2701	Qy	TTCTCACTGGTGTGAATCTTACCTGGCAATCCGGTTCGATGCTACTTCTTCAAACTG
2701	Db	TTCTCACTGGTGTGAATCTTACCTGGCAATCCGGTTCGATGCTACTTCTTCAAACTG	2761	Qy	ATTCAAAATGGATTTAAATATGGATCAGCCATTTTATGCTACAGAAAGAAATATACATTTATA
2761	Db	ATTCAAAATGGATTTAAATATGGATCAGCCATTTTATGCTACAGAAAGAAATATACATTTATA	2821	Qy	ATGCAATCTCTGTGGGCTTTTGAATCAAGGAAACGGGCATCTTTCCTTAATTTGACTTCG
2821	Db	ATGCAATCTCTGTGGGCTTTTGAATCAAGGAAACGGGCATCTTTCCTTAATTTGACTTCG	2881	Qy	TTTTTAGATGTCTGCTGAGATGTCAGAGAGAGCGATGATCAGATGTGGATGGCCGT
2881	Db	TTTTTAGATGTCTGCTGAGATGTCAGAGAGAGCGATGATCAGATGTGGATGGCCGT	2941	Qy	GTTTTCGGAATGTCTCTCACACTACAAAAACAGGGCACTTTTCAGCAGAGAGATGGATTG
2941	Db	GTTTTCGGAATGTCTCTCACACTACAAAAACAGGGCACTTTTCAGCAGAGAGATGGATTG	3001	Qy	GGGCTACTCTCTCCGAGGCTCAGGTCCGAAGATTCGTAGAGGAACAAATTTAGTTTTAAGG
3001	Db	GGGCTACTCTCTCCGAGGCTCAGGTCCGAAGATTCGTAGAGGAACAAATTTAGTTTTAAGG	3061	Qy	ACTTCATCCCTCACCCTATCCCTAACAGGCTGTAGATTACAAAAAAGAATTTAATTTTCAT
3061	Db	ACTTCATCCCTCACCCTATCCCTAACAGGCTGTAGATTACAAAAAAGAATTTAATTTTCAT	3121	Qy	CACATAAAGAAAAATCTATTATCAACTGCTGCTTCCACAGACTTTTCTCTAGAGAGCGTCT
3121	Db	CACATAAAGAAAAATCTATTATCAACTGCTGCTTCCACAGACTTTTCTCTAGAGAGCGTCT	3181	Qy	CGGTTTACTCTGTGTTTTCAAGGAGCACTTTTGTAAAAATCAAAATCATCTCTGTCAGAGGCG
3181	Db	CGGTTTACTCTGTGTTTTCAAGGAGCACTTTTGTAAAAATCAAAATCATCTCTGTCAGAGGCG	3241	Qy	GAGAGCTGATATGAACTTTTATTTGGAGCATTTGATCTGCATCCAAAGGCTTCTCAGGCGG
3241	Db	GAGAGCTGATATGAACTTTTATTTGGAGCATTTGATCTGCATCCAAAGGCTTCTCAGGCGG			

QY	3301	GCCTGAGTGAATGTGTGTAACCTGAAGTACAGTATATCTTCTGTAATACATAAAACAAAAGC	3360
DB	3301	GCCTGAGTGAATGTGTGTAACCTGAAGTACAGTATATCTTCTGTAATACATAAAACAAAAGC	3360
QY	3361	ATTTTGTCTAAGGAGACCTAATATGATTTTTTTTAAGTCTATGTTTTTAAATAATAATATGTAAA	3420
DB	3361	ATTTTGTCTAAGGAGACCTAATATGATTTTTTTTAAGTCTATGTTTTTAAATAATAATATGTAAA	3420
QY	3421	TTTTTTCAGCTATTAGTGATATATTTTTTATCGGTGGGAATAAAATTTCTACTACAG	3475
DB	3421	TTTTTTCAGCTATTAGTGATATATTTTTTATCGGTGGGAATAAAATTTCTACTACAG	3475
RESULT 7			
HSU02687			
LOCUS		3475 bp mRNA linear PRI 11-JUN-1994	
DEFINITION		Human growth factor receptor tyrosine kinase (SYK-1) mRNA, complete cds.	
ACCESSION	U02687		
VERSION	U02687.1	GI:409572	
KEYWORDS			
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
AUTHORS		Small,D., Levenstein,M., Kim,E., Carow,C., Amin,S., Rockwell,P., Witte,L., Burrow,C., Ratajczak,M.Z., Gewirtz,A.M. et al.	
TITLE		SYK-1, the human homolog of Flk-2/Fit-3, is selectively expressed in CD34+ human bone marrow cells and is involved in the proliferation of early progenitor/stem cells	
JOURNAL		Proc. Natl. Acad. Sci. U.S.A. 91 (2), 459-463 (1994)	
MEDLINE	94119906		
PUBMED	7507245		
REFERENCE		2 (bases 1 to 3475)	
AUTHORS		Small,D.	
TITLE		Direct Submission	
JOURNAL		Submitted (25-OCT-1993) Donald Small, Oncology, Johns Hopkins University School of Medicine, 600 N. Wolfe St., Baltimore, MD 21287 USA	
FEATURES		Location/Qualifiers	
source		1. 3475	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/map="13q12"	
		/cell_type="stem cell"	
		/tissue_type="bone marrow"	
		58..3039	
CDS		/codon_start=1	
		/product="serine/threonine protein kinase"	
		/protein_id="AAAI8947.1"	
		/db_xref="GI:409573"	
		/translation="MPALARDAGTPLLVVFSAMIFGTITNODLPVFKICVLINHKND SSVKGSSPVMSSEDPALDGCALRPOSGTVYEAADVDSATIFLOVLVDPAGNIS LWFKVHSLINCPKPHNDLRGVGVSVILKMTETOAGEYLLFQSEATNTYTLFTVSI RLTWLYTLRRPYPRKMNODLVCISVEPIAEVWELCDSGESCKESPAVKKKE KVLUHELFDTIRCCARNEGRECTRFLTDLNTQPTTLPQLFKVGEPLWIRKCAVH VHQSGALVLELKNQIAOVAMILEKEKADSSERALEMSELKQMTQLGSHENIVLLGACT HPQSFALTVKGFNINATNSSEYDIQYEFCSFRKAYPQIRCTWTFRKPSPCK EOKGLONGYSISKFNHQPGEYIPHAENDDAOFKMTFLAIRRKPQVLARASAOA SCFSDGVPSPWTWKCSDKSPNCTEETEGVYNKRNKRVGOWVSSSTIANSEALK GFVLKCAYNLSGTSCEITLLNSPQFPFIQNIISYATIGVCLLFIIVITLLICHKI KQFQRYESQLQWVQVGGSDNIEFYVDPREYIDLKWEFPRENLEPGKVLGSGAGFKY MNATYGIKSTGYQVQVAMILEKEKADSSERALEMSELKQMTQLGSHENIVLLGACT LSGPIYLIPEYCCYGLDNLNLSKREKPHRTWTETPKENHFSYPTTFQSPHPNSMPGS REVOIHDPDSQISGLHNSFHSDETEYENOKLREEDLNVLTPEDLLCFAYOVKMA MEPLEFKSVCHRDLAARNVLVTHGKVVKICDPLGARDIMSDGNYVVRGNARLPYQVMA PESLPFGITVTKISDWSYGLLWEIPSLGVNYPFGIPVDANFYKLQNGFKMDQPPYA TESIYIIMQSCWAFDSRKGKPSFNLTSFLGCOLADAEAMYNQVDRSECPTYONR RPFSEMDLGLLSPQAVEDS"	
BASE COUNT	1042 a	709 g	940 t



## ORIGIN

Query Match 99.3%; Score 3475; DB 9; Length 3475;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3475; Conservative 0; Mismatches 0; Indels 0

QY	1	CGAGGCGGCATCCGAGGCGTGGGCGGCGCCCTGGGGGACCCCGGGCTCCGGAGCCCATG	60
DB	1	CGAGGCGGCATCCGAGGCGTGGGCGGCGCCCTGGGGGACCCCGGGCTCCGGAGCCCATG	60
QY	61	CCGCGGTGGCGCGCAGCGGGGACCGGTGCGCTGCTGCTGTTGTTTCTGCAATGATA	120
DB	61	CCGCGGTGGCGCGCAGCGGGGACCGGTGCGCTGCTGCTGTTGTTTCTGCAATGATA	120
QY	121	TTTGGGACTATTA CAATCAAGATCTGCTGTGTGATCAAGTGTGTTTAAATCAATCAATG	180
DB	121	TTTGGGACTATTA CAATCAAGATCTGCTGTGTGATCAAGTGTGTTTAAATCAATCAATG	180
QY	181	AACAAATGATTCATCAGTGGGGAAGTCATCATATCCCATGGTATCAGAAATCCCGGAA	240
DB	181	AACAAATGATTCATCAGTGGGGAAGTCATCATATCCCATGGTATCAGAAATCCCGGAA	240
QY	241	GACCTCGGGTGTGGTTGAGACCCAGAGCTCAGGACAGTGTACGAAGCTGCCGCTGTG	300
DB	241	GACCTCGGGTGTGGTTGAGACCCAGAGCTCAGGACAGTGTACGAAGCTGCCGCTGTG	300
QY	301	GAAGTGAATGTATCTGCTTCCATCACA CTGCAAGTGTGTGTGATGCCCGAGGAAAT	360
DB	301	GAAGTGAATGTATCTGCTTCCATCACA CTGCAAGTGTGTGTGATGCCCGAGGAAAT	360
QY	361	TCCTGCTCTCGGGTCTTTAAGCACAGCTCCCTGAA TTGCCAGCCACATTTTGATTTTCAA	420
DB	361	TCCTGCTCTCGGGTCTTTAAGCACAGCTCCCTGAA TTGCCAGCCACATTTTGATTTTCAA	420
QY	421	AACAGAGAGTGTGTTTCCATGTGTCA TTTGAAAATGACAGAAACCCAGCTGGAGAAATC	480
DB	421	AACAGAGAGTGTGTTTCCATGTGTCA TTTGAAAATGACAGAAACCCAGCTGGAGAAATC	480
QY	481	CTACTTTTATTCAGAGTGAAGCTACCAATTA CACAATATTTGTTTACAGTGAATGAAGA	540
DB	481	CTACTTTTATTCAGAGTGAAGCTACCAATTA CACAATATTTGTTTACAGTGAATGAAGA	540
QY	541	AATACCTCTGTTTACATTAAGAAGACCTTACTTTAGAAAAATGGA AAAACCCAGGAGCC	600
DB	541	AATACCTCTGTTTACATTAAGAAGACCTTACTTTAGAAAAATGGA AAAACCCAGGAGCC	600
QY	601	CTGGTCTGCATATCTGAGAGCGTTTCCAGAGCCGATCGTGGAA TGGGTGCTTTGCGATTC	660
DB	601	CTGGTCTGCATATCTGAGAGCGTTTCCAGAGCCGATCGTGGAA TGGGTGCTTTGCGATTC	660
QY	661	CAGGGGGAAGCTGTAAAGAAAGATCCAGCTGTTGTTTAAAAGGAGGAAAAGTGCCTT	720
DB	661	CAGGGGGAAGCTGTAAAGAAAGATCCAGCTGTTGTTTAAAAGGAGGAAAAGTGCCTT	720
QY	721	CATGAATTAATTTGGGACGGACATAAAGTGTGTCGCCAGAAATGAACTGGGCGAGGAATGC	780
DB	721	CATGAATTAATTTGGGACGGACATAAAGTGTGTCGCCAGAAATGAACTGGGCGAGGAATGC	780
QY	781	ACCAGGCTGTTTCAATATAGATCTTAAATCAAACTCCTCAGACCAATTTGCCACAAATTA	840
DB	781	ACCAGGCTGTTTCAATATAGATCTTAAATCAAACTCCTCAGACCAATTTGCCACAAATTA	840
QY	841	CTTAAATAGTGGGAAACCTTATGATTAAGTGC AAAAGCTGTTCAATGTGAAACCAATGATTC	900
DB	841	CTTAAATAGTGGGAAACCTTATGATTAAGTGC AAAAGCTGTTCAATGTGAAACCAATGATTC	900
QY	901	GGGCTCACCTGGGAAATTAGAAAAACAAAGCACTCGAGGAGGGCAACTACTTTTGAGATGAGT	960
DB	901	GGGCTCACCTGGGAAATTAGAAAAACAAAGCACTCGAGGAGGGCAACTACTTTTGAGATGAGT	960
QY	961	ACCTATTCAA CAACAGAACTATGATA CGGAATTCGTTTGTCTTTGTATCATCAGTGGCA	1020
DB	961	ACCTATTCAA CAACAGAACTATGATA CGGAATTCGTTTGTCTTTGTATCATCAGTGGCA	1020

Qy	1021	AGAAAGACACCGGATAC	TACTACATGTTGTTCTCTTCAAAGCATCCCAAGTCAATCAGCTTTG	1080
Db	1021	AGAAACACACCGGATAC	TACTACATGTTGTTCTCTTCAAAGCATCCCAAGTCAATCAGCTTTG	1080
Qy	1081	GTTACCATCTGAGGAAGGAT	TATAAATGCTACCAATTCAGTGAAGATATAGAAATT	1140
Db	1081	GTTACCATCTGAGGAAGGAT	TATAAATGCTACCAATTCAGTGAAGATATAGAAATT	1140
Qy	1141	GACCAATATGAAGAGTTT	TTGTTTTCTGTCAGGTTTAAAGCCTTACCCAATAATCAGATGT	1200
Db	1141	GACCAATATGAAGAGTTT	TTGTTTTCTGTCAGGTTTAAAGCCTTACCCAATAATCAGATGT	1200
Qy	1201	ACGTGGACCTTC	TCTCGAAAATCATTTCTCTGTCAGGTTTAAAGCCTTACCCAATAATCAGATGT	1260
Db	1201	ACGTGGACCTTC	TCTCGAAAATCATTTCTCTGTCAGGTTTAAAGCCTTACCCAATAATCAGATGT	1260
Qy	1261	AGCATATCCAAAGTTT	TGCAATCATAGCACCGCAGGAGAAATATATATTCCTCATGCAGAA	1320
Db	1261	AGCATATCCAAAGTTT	TGCAATCATAGCACCGCAGGAGAAATATATATTCCTCATGCAGAA	1320
Qy	1321	AATGATGATGCCAAATTT	TACCAAATGTTTCACTGCTGAAATATAGAAAGGAACCTCAAGTG	1380
Db	1321	AATGATGATGCCAAATTT	TACCAAATGTTTCACTGCTGAAATATAGAAAGGAACCTCAAGTG	1380
Qy	1381	CTCGAAGACATCGGCAAGT	TCAGGCGTCTGTTTCTCGAATGGATACCCATTTACCATCT	1440
Db	1381	CTCGAAGACATCGGCAAGT	TCAGGCGTCTGTTTCTCGAATGGATACCCATTTACCATCT	1440
Qy	1441	TGACCTGGAAAGTGTTT	TCAGACAAAGTCTCCAACTGCAACAGAGATCAAGAAAGGA	1500
Db	1441	TGACCTGGAAAGTGTTT	TCAGACAAAGTCTCCAACTGCAACAGAGATCAAGAAAGGA	1500
Qy	1501	GTCTGGAATAGAAAGCT	TACAGAAAGTGTGTTGGACAGTGGGTGTCGACGATCTCTA	1560
Db	1501	GTCTGGAATAGAAAGCT	TACAGAAAGTGTGTTGGACAGTGGGTGTCGACGATCTCTA	1560
Qy	1561	AAATGATGATGAAAGGCT	TCTGGTCAAAGTGTGTCATCAATTCCTTTGGC	1620
Db	1561	AAATGATGATGAAAGGCT	TCTGGTCAAAGTGTGTCATCAATTCCTTTGGC	1620
Qy	1621	ACATCTTGTGAGACGAT	CTCTTTAAACTCTCAGGCCCCCTTCCCTTTCAATCCAAAGCAAC	1680
Db	1621	ACATCTTGTGAGACGAT	CTCTTTAAACTCTCAGGCCCCCTTCCCTTTCAATCCAAAGCAAC	1680
Qy	1681	ATCTCATCTCATGCAAAAT	TTGGTGTTGTCCTCTTCAATGTCGTTTAAACCTTGCTA	1740
Db	1681	ATCTCATCTCATGCAAAAT	TTGGTGTTGTCCTCTTCAATGTCGTTTAAACCTTGCTA	1740
Qy	1741	ATTGTCACAAAGTACAAAGCAAT	TTAGTATGAAAGCCAGCTACAGATGGTACAGGTG	1800
Db	1741	ATTGTCACAAAGTACAAAGCAAT	TTAGTATGAAAGCCAGCTACAGATGGTACAGGTG	1800
Qy	1801	ACCGGCTCTCAGATAAT	GAGTACTTCTACGTTGATTTTCAGAGAAATGAAATGATCTC	1860
Db	1801	ACCGGCTCTCAGATAAT	GAGTACTTCTACGTTGATTTTCAGAGAAATGAAATGATCTC	1860
Qy	1861	AAATGGAGGTTTCCAGAGAAAT	TTAGAGTTTGGGAAGTACTAGGATCAGGTGCTTTT	1920
Db	1861	AAATGGAGGTTTCCAGAGAAAT	TTAGAGTTTGGGAAGTACTAGGATCAGGTGCTTTT	1920
Qy	1921	GGAAAGTATGAACGCAAC	CAGCTTATGAAATAGCAAAACAGGAGTCTCAATCCAGGTT	1980
Db	1921	GGAAAGTATGAACGCAAC	CAGCTTATGAAATAGCAAAACAGGAGTCTCAATCCAGGTT	1980
Qy	1981	GCGGTCAAATGCTGAAAGAAA	AGCAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAA	2040
Db	1981	GCGGTCAAATGCTGAAAGAAA	AGCAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAA	2040
Qy	2041	CTCAAGATGATGACCCAGTGGG	AAGCCACGAGAAATTTGTGAACTGTGGGGGGCTGC	2100
Db	2041	CTCAAGATGATGACCCAGTGGG	AAGCCACGAGAAATTTGTGAACTGTGGGGGGCTGC	2100



QY 421 AACAGAGGAGTGTGTTCCATGGTCAATTTGAAATATGACAGAAAACCAAGCTGGAGAATAC 480  
Db 421 AACAGAGGAGTGTGTTCCATGGTCAATTTGAAATATGACAGAAAACCAAGCTGGAGAATAC 480  
QY 481 CTACTTTTATTTACAGAGTGAAGCTACCAATTAACAATATGTTTAAAGTGAATATGA 540  
Db 481 CTACTTTTATTTACAGAGTGAAGCTACCAATTAACAATATGTTTAAAGTGAATATGA 540  
QY 541 AATACCTCTGTTTACACATTAAGAAAGACCTTTACTTTAGAAAATGGAATAACAGGAGGCC 600  
Db 541 AATACCTCTGTTTACACATTAAGAAAGACCTTTACTTTAGAAAATGGAATAACAGGAGGCC 600  
QY 601 CTGGTCTGATATCTAGAGCGTTCAGAGCCGATCGTGGAAATGGGTGCTTTCCGATTC 660  
Db 601 CTGGTCTGATATCTAGAGCGTTCAGAGCCGATCGTGGAAATGGGTGCTTTCCGATTC 660  
QY 661 CAGGGGGAAGCTGTAAAGAAAGTCTCAGCTGTTGTTTAAAGAGGAAAGTGCCT 720  
Db 661 CAGGGGGAAGCTGTAAAGAAAGTCTCAGCTGTTGTTTAAAGAGGAAAGTGCCT 720  
QY 721 CATGAATTAATTTGGGACGGACATAAGGTCTGTGCCAGAAATGCACTGGCGCAGGAAATGC 780  
Db 721 CATGAATTAATTTGGGACGGACATAAGGTCTGTGCCAGAAATGCACTGGCGCAGGAAATGC 780  
QY 781 ACCAGGCTGTTCAATATAGATCTAAATCAAACTCCTCAGACCACATTTGCCAATTAATTT 840  
Db 781 ACCAGGCTGTTCAATATAGATCTAAATCAAACTCCTCAGACCACATTTGCCAATTAATTT 840  
QY 841 CTTTAAAGTAGGGAAACCCCTATGATTAAGGTGCAAGCTGTTTCATGTGAACCATGGATTC 900  
Db 841 CTTTAAAGTAGGGAAACCCCTATGATTAAGGTGCAAGCTGTTTCATGTGAACCATGGATTC 900  
QY 901 GGGCTCACCTGGGAATTAAGAAACAAAGCACTCGAGGAGGGCAACTACTTTGAGATGAT 960  
Db 901 GGGCTCACCTGGGAATTAAGAAACAAAGCACTCGAGGAGGGCAACTACTTTGAGATGAT 960  
QY 961 ACCTATTCAACAAACAGAACTATGATACGGATTCCTGTTGCTTTGATATCATCAGTGCA 1020  
Db 961 ACCTATTCAACAAACAGAACTATGATACGGATTCCTGTTGCTTTGATATCATCAGTGCA 1020  
QY 1021 AGAAACGACACCGGATACACTGTTGTTCTTTCAAAAGCATCCAGTCAATCAGCTTTG 1080  
Db 1021 AGAAACGACACCGGATACACTGTTGTTCTTTCAAAAGCATCCAGTCAATCAGCTTTG 1080  
QY 1081 GTTACCATCTAGGAAAGGATTTATAAATGCTACCAATTCAGTGAAGATTAATAAT 1140  
Db 1081 GTTACCATCTAGGAAAGGATTTATAAATGCTACCAATTCAGTGAAGATTAATAAT 1140  
QY 1141 GACCAATATGAAGATTTGTTTCTGTCAGGTTTAAAGCCTACCCACCAATCAGATGT 1200  
Db 1141 GACCAATATGAAGATTTGTTTCTGTCAGGTTTAAAGCCTACCCACCAATCAGATGT 1200  
QY 1201 ACCTGACCTTCTCTCGAAATCAATTTCTTTGTGAGCAAAAGGCTTTGATAACGGATAC 1260  
Db 1201 ACCTGACCTTCTCTCGAAATCAATTTCTTTGTGAGCAAAAGGCTTTGATAACGGATAC 1260  
QY 1261 AGCATATCCAAGTTTTCGAATCATAGCAACAGCAGGAGAGATATATTTCCATGCGAGAA 1320  
Db 1261 AGCATATCCAAGTTTTCGAATCATAGCAACAGCAGGAGAGATATATTTCCATGCGAGAA 1320  
QY 1321 AATGATGATGCCCAATTTTACCAAAATGTTTACCGCTGATATAGAGGAAACCTCAAGTG 1380  
Db 1321 AATGATGATGCCCAATTTTACCAAAATGTTTACCGCTGATATAGAGGAAACCTCAAGTG 1380  
QY 1381 CTCGCAAGAGCATCGGCAAGTCAAGGCTGCTGTTTCTCGATGATACCCATTTACCATCT 1440  
Db 1381 CTCGCAAGAGCATCGGCAAGTCAAGGCTGCTGTTTCTCGATGATACCCATTTACCATCT 1440  
QY 1441 TGGACCTGGAAGAGTGTTCAGAAAGTCTCCCAACTGCAAGAGAGATCAAGAGGA 1500  
Db 1441 TGGACCTGGAAGAGTGTTCAGAAAGTCTCCCAACTGCAAGAGAGATCAAGAGGA 1500  
QY 1501 GTCTGGAATAGAAAGGCTTAAACAGAAAAGTGTGTCAGAGTGGGTGTCGAGCAGTACTCTA 1560

Db 1501 GTCTGGAATAGAAAGGCTTAAACAGAAAAGTGTGTCAGAGTGGGTGTCGAGCAGTACTCTA 1560  
QY 1561 AACATGAGTGAAGCCATTAAGAGGTTCTCGTCAAGTGTGTGTCATACAATTCCTCTGGC 1620  
Db 1561 AACATGAGTGAAGCCATTAAGAGGTTCTCGTCAAGTGTGTGTCATACAATTCCTCTGGC 1620  
QY 1621 ACATCTTGAGAGCGATCCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCAATCCAGACAAC 1680  
Db 1621 ACATCTTGAGAGCGATCCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCAATCCAGACAAC 1680  
QY 1681 ATCTCAATCTATGCAACAAATGTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740  
Db 1681 ATCTCAATCTATGCAACAAATGTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740  
QY 1741 ATTTGTCAACAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
Db 1741 ATTTGTCAACAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
QY 1801 ACCGGCTCTCAGATAATGATCTCTAGCTTGTGTTGGAAGGTACTAGGATCAGGTCTCTT 1860  
Db 1801 ACCGGCTCTCAGATAATGATCTCTAGCTTGTGTTGGAAGGTACTAGGATCAGGTCTCTT 1860  
QY 1861 AAATGGGATTTCCAGAGAAAATTTAGAGTTTGGAGGTACTAGGATCAGGTCTCTT 1920  
Db 1861 AAATGGGATTTCCAGAGAAAATTTAGAGTTTGGAGGTACTAGGATCAGGTCTCTT 1920  
QY 1921 GGAAAGTGTATGAAACGCAACAGCTTATGGAAATAGCAAAACAGGAGTCTCAATCCAGTT 1980  
Db 1921 GGAAAGTGTATGAAACGCAACAGCTTATGGAAATAGCAAAACAGGAGTCTCAATCCAGTT 1980  
QY 1981 GCCGTCAAAATGCTCAAAAGAAAAGCAGACAGCTCTGAAAGAGGCACTCATGTGAGAA 2040  
Db 1981 GCCGTCAAAATGCTCAAAAGAAAAGCAGACAGCTCTGAAAGAGGCACTCATGTGAGAA 2040  
QY 2041 CTCAGATGATGACCCAGCTGGAGCAACGAGAAATATTGAAATCTGCTGGGGGGCTGC 2100  
Db 2041 CTCAGATGATGACCCAGCTGGAGCAACGAGAAATATTGAAATCTGCTGGGGGGCTGC 2100  
QY 2101 ACCTGTGAGGACCAATTTATGATTTTGAATCTGTTGCTATGCTGATCTTCTCAAC 2160  
Db 2101 ACCTGTGAGGACCAATTTATGATTTTGAATCTGTTGCTATGCTGATCTTCTCAAC 2160  
QY 2161 TATCTAAGATTAAGAGAAAATTTCAAGGACTTGGACAGAGATTTCAAGAAACAC 2220  
Db 2161 TATCTAAGATTAAGAGAAAATTTCAAGGACTTGGACAGAGATTTCAAGAAACAC 2220  
QY 2221 AATTTCAAGTTTACCCCACTTTCCAAATCACATCCAAATCCAAATCCAAATCCAAATCCAAAT 2280  
Db 2221 AATTTCAAGTTTACCCCACTTTCCAAATCACATCCAAATCCAAATCCAAATCCAAATCCAAAT 2280  
QY 2281 GAAGTTCAGATACACCCGACTCCGACTCGATCAATCTCAGGGCTTCAATGGAAATTTCAATTCAC 2340  
Db 2281 GAAGTTCAGATACACCCGACTCCGACTCGATCAATCTCAGGGCTTCAATGGAAATTTCAATTCAC 2340  
QY 2341 TCTGAAGATGAATTTGAATATGAAACCAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400  
Db 2341 TCTGAAGATGAATTTGAATATGAAACCAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400  
QY 2401 GTGCTTACATTTGAAAGATCTTTCTGTTTGCATATCAGTTGCTGCAAGAAATTTT 2460  
Db 2401 GTGCTTACATTTGAAAGATCTTTCTGTTTGCATATCAGTTGCTGCAAGAAATTTT 2460  
QY 2461 CTGGAATTTAAGTCTGTTTCAAGAGACCTGGCCGCCAGGAAACGTCCTGTCACCCAC 2520  
Db 2461 CTGGAATTTAAGTCTGTTTCAAGAGACCTGGCCGCCAGGAAACGTCCTGTCACCCAC 2520  
QY 2521 GGGAAAGTGGTGAAGATTTGATGCTTTGGATTTGGCTCGAGATATCATGATGATTTCAAC 2580  
Db 2521 GGGAAAGTGGTGAAGATTTGATGCTTTGGATTTGGCTCGAGATATCATGATGATTTCAAC 2580  
QY 2581 TATGTTGTCAGGGGCAATGCCCTGCTGCTGTAATAATGATGCTGCTGCTGCTGCTGCTGCTGCT 2640

Db 2581 TATGTTGTACGGGCAATGCCCTCTGTCCTGTAATAATGGATGGCCCGCCGAAAGCGCTGTTT 2640

Qy 2641 GAAGGCATCTACACCAATTAAGAGTGATGTCGTGTCATATGAATATTAATCTGTCGGAAATC 2700

Db 2641 GAAGGCATCTACACCAATTAAGAGTGATGTCGTGTCATATGAATATTAATCTGTCGGAAATC 2700

Qy 2701 TTCTCACTGGTGGAATCCTTACCTCGGCAATTCGGGTGATGTCGTAACCTTCTCAAACTG 2760

Db 2701 TTCTCACTGGTGGAATCCTTACCTCGGCAATTCGGGTGATGTCGTAACCTTCTCAAACTG 2760

Qy 2761 ATTCAAAATGGATTAATAATGGATCAGCCATTTATGCTACAGAAATATACATATA 2820

Db 2761 ATTCAAAATGGATTAATAATGGATCAGCCATTTATGCTACAGAAATATACATATA 2820

Qy 2821 ATCAATCTCTCGGCTGTTTGAATCAAGAAACGGCCATCTTCCCTAAATTTGACTTCG 2880

Db 2821 ATCAATCTCTCGGCTGTTTGAATCAAGAAACGGCCATCTTCCCTAAATTTGACTTCG 2880

Qy 2881 TTTTATAGGATGTCAGCTGGGAGATGCAAGAAAGCGATGATACAGAAATGGATGGCGGT 2940

Db 2881 TTTTATAGGATGTCAGCTGGGAGATGCAAGAAAGCGATGATACAGAAATGGATGGCGGT 2940

Qy 2941 GTTTCGGATGTCCTCACACTACCAACCAAGGCGCTTTCACGAGAGATGATTTG 3000

Db 2941 GTTTCGGATGTCCTCACACTACCAACCAAGGCGCTTTCACGAGAGATGATTTG 3000

Qy 3001 GGGCTACTCTCGGAGGCTCAGTGCAGATTCGAGAAACAAATTTAGTTTAAAG 3060

Db 3001 GGGCTACTCTCGGAGGCTCAGTGCAGATTCGAGAAACAAATTTAGTTTAAAG 3060

Qy 3061 ACTTCATCTCTCCACTATCCCTAACAGGCTGTAGATACCAAAACAAAGATTAATTTTAT 3120

Db 3061 ACTTCATCTCTCCACTATCCCTAACAGGCTGTAGATACCAAAACAAAGATTAATTTTAT 3120

Qy 3121 CACTAAAGAAATCTATTAATCACTGCTGCTTCCAGAGATTTTCTTAGAAGCGCTCT 3180

Db 3121 CACTAAAGAAATCTATTAATCACTGCTGCTTCCAGAGATTTTCTTAGAAGCGCTCT 3180

Qy 3181 GCGTTTACTCTGTTTCAAGAGGACTTTTGTAAATCAATCATCTGTCACAGGCGAG 3240

Db 3181 GCGTTTACTCTGTTTCAAGAGGACTTTTGTAAATCAATCATCTGTCACAGGCGAG 3240

Qy 3241 GAGGAGCTGATAATGAATTTATTTGAGCAATGATCTGCATCCAGGCTTCTCAGCGCG 3300

Db 3241 GAGGAGCTGATAATGAATTTATTTGGAGCATTTGATCTGCATCCAGGCTTCTCAGCGCG 3300

Qy 3301 GCTTGAGTGAATTTGTAACCTGAGTACAGTATATCTTGTAAATACATAAACAAGC 3360

Db 3301 GCTTGAGTGAATTTGTAACCTGAGTACAGTATATCTTGTAAATACATAAACAAGC 3360

Qy 3361 ATTTTGCTAAGGAGAGCTAATATGATTTTAAAGTCTATGTTTAAATATATATATAA 3420

Db 3361 ATTTTGCTAAGGAGAGCTAATATGATTTTAAAGTCTATGTTTAAATATATATATAA 3420

Qy 3421 TTTTTCAGCTATTTAGTGATATATTTATTTGGTGGGATAAAATTTCTACTACAGA 3476

Db 3421 TTTTTCAGCTATTTAGTGATATATTTATTTGGTGGGATAAAATTTCTACTACAGA 3476

RESULT 9

HSFLT3RTK

LOCUS HsFLT3RTK 3422 bp mRNA linear PRI 17-JAN-1995

DEFINITION H. sapiens FLT3 mRNA for FLT3 receptor tyrosine kinase.

ACCESSION Z26652 S64785

VERSION Z26652.1 GI:406322

KEYWORDS Flt3 gene; receptor tyrosine kinase.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 3422)

AUTHORS Rosnet, O., Schiff, C., Pebusque, M.J., Marchetto, S., Tonnel, C., Tolron, Y., Birg, F. and Birnbaum, D.

TITLE Human FLT3/FLK2 gene: cDNA cloning and expression in hematopoietic cells

JOURNAL Blood 82 (4), 1110-1119 (1993)

MEDLINE 93357464

PUBMED 8394751

REFERENCE 2 (bases 1 to 3422)

AUTHORS Rosnet, O.

TITLE Direct Submission

JOURNAL Submitted (04-OCT-1993) Rosnet O., Inseem U119, 27, bd. Lei Roure, Marseille, France

COMMENT data in S64785.

FEATURES

Location/Qualifiers

1..3422

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

1..2982

/gene="FLT3"

1..2982

/gene="FLT3"

/codon\_start=1

/product="FLT3 receptor tyrosine kinase"

/protein\_id="CAA81393.1"

/db\_xref="GI:406323"

/db\_xref="SWISS-PROT:P36888"

/translation="MPALARDGGQLPVLVVSAMIFGTITNQDLPLVIKCVLINHKNND

SSVKSSTSYPMVSESPEDLGCALRPOSSGTVYERAAVEVDVASITLQVLVDAPGNIS

CLWVFKHSLNCPKFDLQNRGVSMVILKMTQAGEYLLFQSEATNTYILFTVSI

RNTLLYTLRRPYKRMENODALVCISSEVPEIWEVLCDQSGESKEESPAVVKKE

KVLHELPGMDIRCCARNELGRECTRLTIDNQTPOTLTLPQLKVGELWIRCKAVH

VNHPGGLTWELENKALEEGNYFEMSTYSNTRMLRILPAFVSSVARNDGTGYTCSSSK

HPQSALVTIVEKGIPOINATSSDEYDQYEFCSVRFKAYPOIRCTWTFSRKSPFC

EQKGLDNGYSISKFNHKGQGYIIFHAENDDAQFTKMTLINRRKQVLAASQA

SCPSDGVPLSPMTWKCKSPNCTBEITGVWNRKANRVFGOWSSSTILNMEAIK

GLVAKCCAYNSLCTCETILLNSGPPPIQDNISFYATIGVCLLFIWLTLLICHKY

KKQFRESQLQWQVTSNEDYFVDPREYEDLKEWPRENLKPKVLSGAGFKV

MNATATYISKITGVSIOQAVMLKEKADSSERELMSELMQTLGSHENIVNLGACT

LSGPYLIPEYCCYGLLNLRSKREKPHRTWEIPEKHNFSPYPTTQSHNSMPGS

REVQIHPSDQISGLHNSPHSEDEIYENQKLEEBEDLNVLTFELLCPAYQVAKG

MEFLPKSCVHRDLAARNVLTHGVKVICDPLGLARDIMSDSNVVRGNARLPVKWMA

PESLFGIYVTKSDVMSYGLLWEIIFSLGVNYPGIPVDANFYKLIQNGFMDQPFYA

TEBIYIIMQSCWAFDSKRKPSFPNLISFLCQLADAEAMYQNDVGRVSECPHYQNR

RPFRESMDLGLLSLPQAQVEDS"

sig\_peptide 1..69

/gene="FLT3"

mat\_peptide 70..1980

/gene="FLT3"

/product="FLT3 receptor tyrosine kinase"

2983..3422

polyA\_signal 3401..3405

BASE\_COUNT 1038 a 687 c 758 g 939 t

ORIGIN

Query Match 96.9%; Score 3393.8; DB 9; Length 3422;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 3401; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 58 ATGCCGCGTGTGGCGCGACGGCGCACCGTGCCTGCTGCTGTTGTTTCTGCAATG 117

Db 1 ATGCCGCGTGTGGCGCGACGGCGCACCGTGCCTGCTGCTGTTGTTTCTGCAATG 60

Qy 118 ATATTTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATCAT 177

Db 61 ATATTTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATCAT 120

Qy 178 AAGAACAAATGATTCATCAGTGGGGAAGTCAATCATATCCATGCTATCAGATCCCGG 237

Db 121 AAGAACAAATGATTCATCAGTGGGGAAGTCAATCATATCCATGCTATCAGATCCCGG 180

Qy 238 GAAGACCTCGGCTGTGGTGTAGACCCAGAGCTCAGGACAGTGTACGAAGCTCCCGT 297

Db 181 GAAGACCTCGGCTGTGGTGTAGACCCAGAGCTCAGGACAGTGTACGAAGCTCCCGT 240

Qy	298	GTGGAAGTGGATGATCTGCTTCCATCACACTGCAAGTGGTGGTCGATGCCCGAGGAAC	357	Db	1321	GTGCTCGCAGAAGCAATCGCAAGTCAGGCGTCTCTTTTCGGA TGGATACCAATTAACCA	1380
Db	241	GTGGAAGTGGATGATCTGCTTCCATCACTGCAAGTGGTGGTCGATGCCCGAGGAAC	300	Qy	1438	TCTTGGACCTCGAAGAGTGTTCAGACAAAGTCTCCAACTGCAACAGAGATACAGAA	1497
Qy	358	ATTTCTCTGCTCTGGGCTTTAAGCACAGCTCCCTGAATGCCAGCCACATTTTGATTTA	417	Db	1381	TCTTGGACCTCGAAGAGTGTTCAGACAAAGTCTCCAACTGCAACAGAGATACAGAA	1440
Db	301	ATTTCTCTGCTCTGGGCTTTAAGCACAGCTCCCTGAATGCCAGCCACATTTTGATTTA	360	Qy	1498	GGAGTCTGGAATAGAAAGGCTAAACAGAAAGTGTGTTGGACAGTGGGTGTCGAGCAGTACT	1557
Qy	418	CAAAACAGAGAGTGTGTTTCCATGCTGATTTTGAATAATGACAGAAACCCAGCTGGAGAA	477	Db	1441	GGAGTCTGGAATAGAAAGGCTAAACAGAAAGTGTGTTGGACAGTGGGTGTCGAGCAGTACT	1500
Db	361	CAAAACAGAGAGTGTGTTTCCATGCTGATTTTGAATAATGACAGAAACCCAGCTGGAGAA	420	Qy	1558	CTAAACATGAGTGAAGCCATAAAAGGGTTCCTGGTCAAGTGTGTGCATACATAATTCCTTT	1617
Qy	478	TACCTACTTTTTATTTCAGAGTGAAGCTACCAATTAACAATAATTTTACAGTGAAGTATA	537	Db	1501	CTAAACATGAGTGAAGCCATAAAAGGGTTCCTGGTCAAGTGTGTGCATACATAATTCCTTT	1560
Db	421	TACCTACTTTTTATTTCAGAGTGAAGCTACCAATTAACAATAATTTTACAGTGAAGTATA	480	Qy	1618	GGCACATCTGTGAGACGATCCTTTTTTAAACCTCTCAGGCCCCCTCCCTTTTCATCCAAAGAC	1677
Qy	538	AGAAATACCCCTGCTTTACACATTAAGAGACCTTACTTTTGAATAATGGAAACACAGGAC	597	Db	1561	GGCACATCTGTGAGACGATCCTTTTTTAAACCTCTCAGGCCCCCTCCCTTTTCATCCAAAGAC	1620
Db	481	AGAAATACCCCTGCTTTACACATTAAGAGACCTTACTTTTGAATAATGGAAACACAGGAC	540	Qy	1678	AACATCTCAATTCATGCAACCAATTTGGTGTGTGTCTCTCTTCATTTGTCTGTTTTAACCCCTG	1737
Qy	598	GCCCTGTGTCGATATCTGAGAGCGTTCCAGAGCCGATCGTGGAAATGGGTGCTTTGCGAT	657	Db	1621	AACATCTCAATTCATGCAACCAATTTGGTGTGTGTCTCTCTTCATTTGTCTGTTTTAACCCCTG	1680
Db	541	GCCCTGTGTCGATATCTGAGAGCGTTCCAGAGCCGATCGTGGAAATGGGTGCTTTGCGAT	600	Qy	1738	CTAATTTCTGCAAGATACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAG	1797
Qy	658	TCACAGGGGGAAAGCTGTAAAGAGAAAGTCCAGCTGTTGTTTAAAGAGGAGAAAGTG	717	Db	1691	CTAATTTCTGCAAGATACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAG	1740
Db	601	TCACAGGGGGAAAGCTGTAAAGAGAAAGTCCAGCTGTTGTTTAAAGAGGAGAAAGTG	660	Qy	1798	GTGACCCGCTCCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGAAATATGAAATATGAT	1857
Qy	718	CTTCATGAAATTAATTTGGGACCGGACATAAGGTGCTGTGCCAGAAATGAACTGGCGAGGGAA	777	Db	1741	GTGACCCGCTCCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGAAATATGAAATATGAT	1800
Db	661	CTTCATGAAATTAATTTGGGACCGGACATAAGGTGCTGTGCCAGAAATGAACTGGCGAGGGAA	720	Qy	1858	CTCAATGGGAGTTTCCAAAGAGAAATTTAGAGTTTGGGAGGTACTAGGATCAGGTGCT	1917
Qy	778	TGCACACAGGCTGTTCAAAATAGATCTAAATCAAACCTCCTCAGACCAATGTCACCAATTA	837	Db	1801	CTCAATGGGAGTTTCCAAAGAGAAATTTAGAGTTTGGGAGGTACTAGGATCAGGTGCT	1860
Db	721	TGCACACAGGCTGTTCAAAATAGATCTAAATCAAACCTCCTCAGACCAATGTCACCAATTA	780	Qy	1918	TTTGGAAAGTGATGAAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAG	1977
Qy	838	TTTCTTAAAGTAGGGACCCCTTATGATTAAGTGCAGAGCTGTTTCATGTGACCATGGA	897	Db	1861	TTTGGAAAGTGATGAAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAG	1920
Db	781	TTTCTTAAAGTAGGGACCCCTTATGATTAAGTGCAGAGCTGTTTCATGTGACCATGGA	840	Qy	1978	GTTCGCGTCAAAATGCTGAAAGAGAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCA	2037
Qy	898	TTTCGGGCTCACCTGGGAATTAGAAAACAAAGCACCTCGAGGAGGGCAACTACTTTTCAGATG	957	Db	1921	GTTCGCGTCAAAATGCTGAAAGAGAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCA	1980
Db	841	TTTCGGGCTCACCTGGGAATTAGAAAACAAAGCACCTCGAGGAGGGCAACTACTTTTCAGATG	900	Qy	2038	GAACTCAAGATGATGACCCAGCTGGGAAAGCCACAGAAATATTTGTGAACTGTCTGGGGCG	2097
Qy	958	AGTACCTATTCAACAAAACAGAACTATGATACGGATTCTGTTGCTTTTGTATCATCAGTG	1017	Db	1981	GAACTCAAGATGATGACCCAGCTGGGAAAGCCACAGAAATATTTGTGAACTGTCTGGGGCG	2040
Db	901	AGTACCTATTCAACAAAACAGAACTATGATACGGATTCTGTTGCTTTTGTATCATCAGTG	960	Qy	2098	TGCACACTGTGAGGACCAATTTTACTTTGATTTTGTGATTTGCTATGCTGATCTTCTC	2157
Qy	1018	GCAAGAAAACGACACCGGATACACTGTTGCTCTTCAAGGATCCGACGTCATCAGCT	1077	Db	2041	TGCACACTGTGAGGACCAATTTTACTTTGATTTTGTGATTTGCTATGCTGATCTTCTC	2100
Db	961	GCAAGAAAACGACACCGGATACACTGTTGCTCTTCAAGGATCCGACGTCATCAGCT	1020	Qy	2158	AACATCTCAAGAGTAAAGAGAGAAATTTTCAGGAGCTTTGGACAGAGATTTTTCAGGAA	2217
Qy	1078	TTGGTTACCATCGTAGGAAAGGATTTTAAATGCTACCAATTCAGGTGAAGATTATGAA	1137	Db	2101	AACATCTCAAGAGTAAAGAGAGAAATTTTCAGGAGCTTTGGACAGAGATTTTTCAGGAA	2160
Db	1021	TTGGTTACCATCGTAGGAAAGGATTTTAAATGCTACCAATTCAGGTGAAGATTATGAA	1080	Qy	2218	CACAAATTTTCAATTTTACCCCACTTTTCCAAATTCACATCCAAATTTCCAGCATGCTGGTTCA	2277
Qy	1138	ATTGACCAATATGAAGAGTTTGTGTTTCTGTGAGTTTAAAGCCCTACCCCAAAATCAGA	1197	Db	2161	CACAAATTTTCAATTTTACCCCACTTTTCCAAATTCACATCCAAATTTCCAGCATGCTGGTTCA	2220
Db	1081	ATTGACCAATATGAAGAGTTTGTGTTTCTGTGAGTTTAAAGCCCTACCCCAAAATCAGA	1140	Qy	2278	AGAGAGTTTCAAGATACCCCGGATCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTT	2337
Qy	1198	TGTACGTGGACCTCTCTCGAAAATCATTTCTGTGAGGAAAAGGGTCTTGATTAACGGA	1257	Db	2221	AGAGAGTTTCAAGATACCCCGGATCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTT	2280
Db	1141	TGTACGTGGACCTCTCTCGAAAATCATTTCTGTGAGGAAAAGGGTCTTGATTAACGGA	1200	Qy	2338	CACCTCTGAAGATGAAATTTGAATAATGAAACCCAAAGAGCTGGAAGAGAGAGGAGACTTG	2397
Qy	1258	TACAGCATATCAAGTTTTCGAAATCATAAGACACAGCCAGGAGATATATATTCATGCA	1317	Db	2281	CACCTCTGAAGATGAAATTTGAATAATGAAACCCAAAGAGCTGGAAGAGAGAGGAGACTTG	2340
Db	1201	TACAGCATATCAAGTTTTCGAAATCATAAGACACAGCCAGGAGATATATATTCATGCA	1260	Qy	2398	AATGTGCTTACATTTGAAGATCTTCTTTGCTTTGCATATCAAGTTGCCAAAGGAAATGGAA	2457
Qy	1318	GAAATATGATGATGCCCAATTTACCAAAATGTTTCAGCTGAAATATGAAGAGGAAACCTCAA	1377	Db	2341	AATGTGCTTACATTTGAAGATCTTCTTTGCTTTGCATATCAAGTTGCCAAAGGAAATGGAA	2400
Db	1261	GAAATATGATGATGCCCAATTTACCAAAATGTTTCAGCTGAAATATGAAGAGGAAACCTCAA	1320	Qy	2458	TTTCTGGAATTAAGTGTGTGTTTTCAGAGACCTGGCGGCCAGGAACTGTGCTTGTCCACC	2517
Qy	1378	GTGCTCGCAGAAGCAATCGGAGTCAAGGCGTCTGTTTCTCGGATGGATACCAATTAACCA	1437				





Dd	781	TTTCTTAAAGTAGGGAAACCCCTTATGGATAAGGTGCAAAAGCTGTTTCATGTGAACCATGGA	840
Qy	898	TTCCGGCTCACCTGGGAATTAGAAAACAAAGACATCGAGGAGGCAACTCTTTGAGATG	957
Dd	841	TTCCGGCTCACCTGGGAATTAGAAAACAAAGACATCGAGGAGGCAACTCTTTGAGATG	900
Qy	958	AGTACCTATTCAAACAAACAGAACTATGATACGGATTCTGTTTGTCTTTGTTGATCATCAGT	1017
Dd	901	AGTACCTATTCAAACAAACAGAACTATGATACGGATTCTGTTTGTCTTTGTTGATCATCAGT	960
Qy	1018	GCAAGAAACGACACCGGATCTACATTTGTTCTCTTCAAAAGCATCCAGTCAATCACT	1077
Dd	961	GCAAGAAACGACACCGGATCTACATTTGTTCTCTTCAAAAGCATCCAGTCAATCACT	1020
Qy	1078	TTGGTTACCATCGTAGGAAGGATTTATAAATGCTACCAATTAAGTGAAGATATGAA	1137
Dd	1021	TTGGTTACCATCGTAGGAAGGATTTATAAATGCTACCAATTAAGTGAAGATATGAA	1080
Qy	1138	ATTGACCAATATGAAGAGTTTGTGTTTCTGTCAGGTTTAAGGCTACCCCAAAATCAGA	1197
Dd	1081	ATTGACCAATATGAAGAGTTTGTGTTTCTGTCAGGTTTAAGGCTACCCCAAAATCAGA	1140
Qy	1198	TGTACGTGGACCTTCTCTCGAAAATCAATTCCTTGTGAGCAAAAAGGCTTTGATAACGGA	1257
Dd	1141	TGTACGTGGACCTTCTCTCGAAAATCAATTCCTTGTGAGCAAAAAGGCTTTGATAACGGA	1200
Qy	1258	TACAGCATATCCAAAGTTTGGCAATCAATGACACAGCAGAGAGATATATATCCATGCA	1317
Dd	1201	TACAGCATATCCAAAGTTTGGCAATCAATGACACAGCAGAGAGATATATATCCATGCA	1260
Qy	1318	GAAATCATGATGCCCAATTTACCAAAATGTTTACGGCTGATATTAAGAGAAACCTCAA	1377
Dd	1261	GAAATCATGATGCCCAATTTACCAAAATGTTTACGGCTGATATTAAGAGAAACCTCAA	1320
Qy	1378	GTGCTCGCAGAAGCATCGCAAGTCAGCGCTGTTTCTCGGATGATACCCATTACCA	1437
Dd	1321	GTGCTCGCAGAAGCATCGCAAGTCAGCGCTGTTTCTCGGATGATACCCATTACCA	1380
Qy	1438	TCCTGGACCTGGAGAGTGTTCAGACAAGTCTCCCAACTGCACAGAAAGAGATCACAGAA	1497
Dd	1381	TCCTGGACCTGGAGAGTGTTCAGACAAGTCTCCCAACTGCACAGAAAGAGATCACAGAA	1440
Qy	1498	GGAGTCGGAATAGAAAGGCTTAACAAAATGTTTACGGCTGATATTAAGAGAAACCTCAA	1557
Dd	1441	GGAGTCGGAATAGAAAGGCTTAACAAAATGTTTACGGCTGATATTAAGAGAAACCTCAA	1500
Qy	1558	CTAAACATGATGATGAAGCCATTAAGAGGTTCTGTCGTCAGTGTGTCGATACAAATCCCTT	1617
Dd	1501	CTAAACATGATGATGAAGCCATTAAGAGGTTCTGTCGTCAGTGTGTCGATACAAATCCCTT	1560
Qy	1618	GGCACAATCTTGTGAGACGATCCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCAATCCAGAC	1677
Dd	1561	GGCACAATCTTGTGAGACGATCCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCAATCCAGAC	1620
Qy	1678	AACATCTCATCTATGCAACAAATGGTGTGTTGTCCTCTTCAATGTCGTTTAAACCCCTG	1737
Dd	1621	AACATCTCATCTATGCAACAAATGGTGTGTTGTCCTCTTCAATGTCGTTTAAACCCCTG	1680
Qy	1738	CTAAATTTGTACAAAGTACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAG	1797
Dd	1681	CTAAATTTGTACAAAGTACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAG	1740
Qy	1798	GTGACCGGCTCTCAGATAATGAGTACTCTAAGTTGATTTTCAAGAAATATGAATATGAT	1857
Dd	1741	GTGACCGGCTCTCAGATAATGAGTACTCTAAGTTGATTTTCAAGAAATATGAATATGAT	1800
Qy	1858	CTCAAAATGGAGTTTCAAGAGAAAATTTAGAGTTTGGGAAGTACTTAGGATCAGGTGCT	1917
Dd	1801	GTCAAAATGGAGTTTCAAGAGAAAATTTAGAGTTTGGGAAGTACTTAGGATCAGGTGCT	1860
Qy	1918	TTTGGAAATGATGAACCCACAGCTTATGGAATTTAGCAAAACAGAGTCTCAATCCAG	1977
Dd	1861	TTTGGAAATGATGAACCCACAGCTTATGGAATTTAGCAAAACAGAGTCTCAATCCAG	1920
Qy	1978	GTTCCCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACCTCATGTCA	2037
Dd	1921	GTTACCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACCTCATGTCA	1980
Qy	2038	GAACTCAAGATGATGACCCAGCTGGAGAGCCACAGAGAAATATGTAACCTGCTGGGGGG	2097
Dd	1981	GAACTCAAGATGATGACCCAGCTGGAGAGCCACAGAGAAATATGTAACCTGCTGGGGGG	2040
Qy	2098	TGCACACTGTGAGACCAATTTACTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGA	2157
Dd	2041	TGCACACTGTGAGACCAATTTACTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGA	2100
Qy	2158	AACATCTAAGAGTAAAGAGAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAA	2217
Dd	2101	AACATCTAAGAGTAAAGAGAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAA	2160
Qy	2218	CACAAATTTCAAGTTTATCCCACTTTTCCAAATCAATCCAAATTTCCAGCATGCTGGTTCA	2277
Dd	2161	CACAAATTTCAAGTTTATCCCACTTTTCCAAATCAATCCAAATTTCCAGCATGCTGGTTCA	2220
Qy	2278	AGAGAAGTTCAGATACACCCGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTT	2337
Dd	2221	AGAGAAGTTCAGATACACCCGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTT	2280
Qy	2338	CACCTCAAGATGAAATTTGAATATGAAAACCAAAAAGGCTGGAAAGAGAGAGGACTTG	2397
Dd	2281	CACCTCAAGATGAAATTTGAATATGAAAACCAAAAAGGCTGGAAAGAGAGAGGACTTG	2340
Qy	2398	AATGCTTACATTTGAAGATCTTCTTGTGCTATCAATCAATCAATCAATCAATCAATCAAT	2457
Dd	2341	AATGCTTACATTTGAAGATCTTCTTGTGCTATCAATCAATCAATCAATCAATCAATCAAT	2400
Qy	2458	TTTCTGGAATTTAAGTCGTGTTTCAAGAGACTGCGCCGAGGAAAGCTGCTGTGCTACC	2517
Dd	2401	TTTCTGGAATTTAAGTCGTGTTTCAAGAGACTGCGCCGAGGAAAGCTGCTGTGCTACC	2460
Qy	2518	CACGGGAAAGTGTGAAGATATGTCGACTTTGGATTGGCTCGAGATATCAATGAGTGTATCC	2577
Dd	2461	CACGGGAAAGTGTGAAGATATGTCGACTTTGGATTGGCTCGAGATATCAATGAGTGTATCC	2520
Qy	2578	AATATGTTCTCAGGGGCAATGCCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2637
Dd	2521	AATATGTTCTCAGGGGCAATGCCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2580
Qy	2638	TTTGAAGGCACTACACCATTAAGAGTGTCTGTCATATGGAATATTAATCTGTGGAA	2697
Dd	2581	TTTGAAGGCACTACACCATTAAGAGTGTCTGTCATATGGAATATTAATCTGTGGAA	2640
Qy	2698	ATCTTCTCACTTGTGTAATCCCTTACCTGCGCATTCGGTTGATGCTGCTGCTGCTGCTG	2757
Dd	2641	ATCTTCTCACTTGTGTAATCCCTTACCTGCGCATTCGGTTGATGCTGCTGCTGCTGCTG	2700
Qy	2758	CTGATTTCAAAATGGAATTTAAATGGAATGATGATGATGATGATGATGATGATGATGATG	2817
Dd	2701	CTGATTTCAAAATGGAATTTAAATGGAATGATGATGATGATGATGATGATGATGATGATG	2760
Qy	2818	ATAAATGCAATCTGCTGGGCTTTTGAATCAAGGAAACGGCCATCTCTTCCCTAAATTTG	2877
Dd	2761	ATAAATGCAATCTGCTGGGCTTTTGAATCAAGGAAACGGCCATCTCTTCCCTAAATTTG	2820
Qy	2878	TGCTTTTATAGGATGTCAGCTGGCAGATGCAAGAGAGGATGATGATGATGATGATGATG	2937
Dd	2821	TGCTTTTATAGGATGTCAGCTGGCAGATGCAAGAGAGGATGATGATGATGATGATGATG	2880
Qy	2938	CGTGTTCGGAATGTCCTTCAACCTTACCAAAACAGGCGGACCTTTTCAGCAGAGATG	2997
Dd	2881	CGTGTTCGGAATGTCCTTCAACCTTACCAAAACAGGCGGACCTTTTCAGCAGAGATG	2940
Qy	2998	TTGGGGCTACTCTCTCCGAGGCTCAGGTGCAAGATTCGTAGAGGAAACAAATTTAGTTT	3057
Dd	2941	TTGGGGCTACTCTCTCCGAGGCTCAGGTGCAAGATTCGTAGAGGAAACAAATTTAGTTT	3000



QY 3058 AGGACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACAAAGAAAGGATTAATTT 3117  
Db 3001 AGGACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACAAAGAAAGGATTAATTT 3060  
QY 3118 CATCACTAAAGAAATCTATTATCACTCTGCTCTCACCAGACTTTCTCTAGAGCG 3177  
Db 3061 CATCACTAAAGAAATCTATTATCACTCTGCTCTCACCAGACTTTCTCTAGAGCG 3120

RESULT 11  
LOCUS I44518 3120 bp DNA linear PAT 07-OCT-1997  
DEFINITION Sequence 22 from patent US 5635177.  
ACCESSION I44518  
VERSION I44518.1 GI:2469231  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 3120)  
AUTHORS Bennett,B.D., Goeddel,D. and Matthews,W.  
TITLE Protein tyrosine kinase agonist antibodies  
JOURNAL Patent: US 5635177-A 22 03-JUN-1997;  
FEATURES  
Location/Qualifiers  
1..3120  
/organism="unknown"  
BASE COUNT 944 a 642 c 700 g 834 t  
ORIGIN

Query Match 88.1%; Score 3083.2; DB 6; Length 3120;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 3097; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 58 ATGCGCGGTGCGCGCGGACGCGGCGACCGTGCCTGCTCGTGTGTTTCTGCAATG 117  
Db 1 ATGAGAGCGTGGCGCGGACGCGGCGGCGGCGTGCCTGCTGTGTTTCTGCAATG 60  
QY 118 ATATTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTAAATCAATCAT 177  
Db 61 ATATTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTAAATCAATCAT 120  
QY 178 AAGAAATGATTCATCAGTGGGGAAGTCAATCATATATCCATGGTATCAGAAATCCCG 237  
Db 121 AAGAAATGATTCATCAGTGGGGAAGTCAATCATATATCCATGGTATCAGAAATCCCG 180  
QY 238 GAAGACTCGGGTGGTGGTGGACCCGAGCTCAGGACAGTGTACGAAGCTGCGCT 297  
Db 181 GAAGACTCGGGTGGTGGTGGACCCGAGCTCAGGACAGTGTACGAAGCTGCGCT 240  
QY 298 GTGGAAGTGGATGTATCTGCTTCCATCACACTGCAAGTGTGGTGGATGCCAGGGAAC 357  
Db 241 GTGGAAGTGGATGTATCTGCTTCCATCACACTGCAAGTGTGGTGGATGCCAGGGAAC 300  
QY 358 ATTTCTGCTCTGGGTGTTTAAAGCAGCTCCCTGAAATGCCAGGCACATTTGATTTA 417  
Db 301 ATTTCTGCTCTGGGTGTTTAAAGCAGCTCCCTGAAATGCCAGGCACATTTGATTTA 360  
QY 418 CAAACAGAGGAGTGTGTTCCATGGTCAATTTGAAATGACAGAAACCCAGCTGGGAA 477  
Db 361 CAAACAGAGGAGTGTGTTCCATGGTCAATTTGAAATGACAGAAACCCAGCTGGGAA 420  
QY 478 TACCTACTTTTATTCAGAGTGAAGCTTACCAATTTACAAATATTTTTCAGTGAATTA 537  
Db 421 TACCTACTTTTATTCAGAGTGAAGCTTACCAATTTACAAATATTTTTCAGTGAATTA 480  
QY 538 AGAAATACCTGCTTTACATTAAGAGACCTTACTTTAGAAATATGGAACCCAGGAC 597  
Db 481 AGAAATACCTGCTTTACATTAAGAGACCTTACTTTAGAAATATGGAACCCAGGAC 540  
QY 598 GGCCTGGTGTGCATATCTGAGAGGTTTCCAGAGCCGATCGTGGAAATGGGTGCTTTGGCAT 657  
Db 541 GGCCTGGTGTGCATATCTGAGAGGTTTCCAGAGCCGATCGTGGAAATGGGTGCTTTGGCAT 600

QY 658 TCACAGGGGAAAGCTGTAAAGAAAGAAAGTCCAGCTGTGTTGTTTAAAGAGGAGGAAAAAGTG 717  
Db 601 TCACAGGGGAAAGCTGTAAAGAAAGAAAGTCCAGCTGTGTTGTTTAAAGAGGAGGAAAAAGTG 660  
QY 718 CTTTCATGAATTTATTCGGGACGGACATAAGGTGCTGTGCCAGAAATGAATGAGCTGGGACGGAA 777  
Db 661 CTTTCATGAATTTATTCGGGACGGACATAAGGTGCTGTGCCAGAAATGAATGAGCTGGGACGGAA 720  
QY 778 TGCACCAAGGCTGTTCAATAGATCTAAATCAAACTCCTCAGACCACATTTGCCCAATTA 837  
Db 721 TGCACCAAGGCTGTTCAATAGATCTAAATCAAACTCCTCAGACCACATTTGCCCAATTA 780  
QY 838 TTTCTTAAAGTAGGGGAAACCTTTATGATGAAGGTGCAAGCTGTTTCATGTGAACCATGGA 897  
Db 781 TTTCTTAAAGTAGGGGAAACCTTTATGATGAAGGTGCAAGCTGTTTCATGTGAACCATGGA 840  
QY 898 TTCGGGCTCACCTGGGAAATTTAGAAACAAAGCACTCGAGGAGGCACTACTTTTGAGATG 957  
Db 841 TTCGGGCTCACCTGGGAAATTTAGAAACAAAGCACTCGAGGAGGCACTACTTTTGAGATG 900  
QY 958 AGTACCTATTTCACAAACAGAACTATGATACGGATCTGTTGCTTTGTTGTTATCATCATG 1017  
Db 901 AGTACCTATTTCACAAACAGAACTATGATACGGATCTGTTGCTTTGTTGTTATCATCATG 960  
QY 1018 GCAAGAAACGACACCGGATACATCATGTTCTCTTCAAAGCATCCAGTCAATCAGCT 1077  
Db 961 GCAAGAAACGACACCGGATACATCATGTTCTCTTCAAAGCATCCAGTCAATCAGCT 1020  
QY 1078 TTGTTTACCATCGTAGGAAAGGATTTATAAATGCTTACCAATTTCAAGTGAAGATTATGAA 1137  
Db 1021 TTGTTTACCATCGTAGGAAAGGATTTATAAATGCTTACCAATTTCAAGTGAAGATTATGAA 1080  
QY 1138 ATTGACCAATATGAAGATTTGTTTCTGTCAGGTTTAAAGCTTACCACCAATCAGA 1197  
Db 1081 ATTGACCAATATGAAGATTTGTTTCTGTCAGGTTTAAAGCTTACCACCAATCAGA 1140  
QY 1198 TGTACGTGGACCTTCTCTGAAAATCATTTCTTGTGAGCAAAAGGCTTTGATAACGA 1257  
Db 1141 TGTACGTGGACCTTCTCTGAAAATCATTTCTTGTGAGCAAAAGGCTTTGATAACGA 1200  
QY 1258 TACAGATATCCAAAGTTTGGCAATCAATGACACCGCAGGAGNAATATATATTCATGCA 1317  
Db 1201 TACAGATATCCAAAGTTTGGCAATCAATGACACCGCAGGAGNAATATATATTCATGCA 1260  
QY 1318 GAAAAATGATGATGCCCAATTTTACCAAAATGTTTCAACGCTGAATATAAGAAAGAAACCTCAA 1377  
Db 1261 GAAAAATGATGATGCCCAATTTTACCAAAATGTTTCAACGCTGAATATAAGAAAGAAACCTCAA 1320  
QY 1378 GTGCTCGAGAAAGCATCGGCAAGTCAAGGCTCTCTGTTTCTCGGATGGATACCCATTAACA 1437  
Db 1321 GTGCTCGAGAAAGCATCGGCAAGTCAAGGCTCTCTGTTTCTCGGATGGATACCCATTAACA 1380  
QY 1438 TCTTGGACCTGGAAGAGTGTTCAGCAAGTCTCCAACTGCAAGAGAGATCACAGAA 1497  
Db 1381 TCTTGGACCTGGAAGAGTGTTCAGCAAGTCTCCAACTGCAAGAGAGATCACAGAA 1440  
QY 1498 GGAAGTCTGGAATAGAAGGCTTACAGAAAGTGTGTTGGAAGTGGGTGTCGAGCAGTACT 1557  
Db 1441 GGAAGTCTGGAATAGAAGGCTTACAGAAAGTGTGTTGGAAGTGGGTGTCGAGCAGTACT 1500  
QY 1558 CTAAACATGATGAAGCCATAAAAAGGTTCTCGGTCAAGTGTGTCATACAAATTCCTCT 1617  
Db 1501 CTAAACATGATGAAGCCATAAAAAGGTTCTCGGTCAAGTGTGTCATACAAATTCCTCT 1560  
QY 1618 GGCACATCTTGTGAGACGATCTTTTAACTCTCCAGGCCCTTCCCTTTTCATCCAGAC 1677  
Db 1561 GGCACATCTTGTGAGACGATCTTTTAACTCTCCAGGCCCTTCCCTTTTCATCCAGAC 1620  
QY 1678 AACATCTCATCTTCTATGCAAAATTTGGGTGTTGTTGCTCTCTTCTTCTGCTTTTAAACCTG 1737  
Db 1621 AACATCTCATCTTCTATGCAAAATTTGGGTGTTGTTGCTCTCTTCTTCTTCTTCTTAAACCTG 1680  
QY 1738 CTAAATTTGTCAAGTACAAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGCTACAG 1797

```
Db      1681  CTAATTGTGTCACAAAGTAAAGCAATTTAGGTATGAAGCCAGCTACAGATGATACAG 1740
Qy      1798  GTACACGGCTCCTCAGATATGAGTACTTCTAAGTTGATTTTCAGAAATATGAATATGAT 1857
Db      1741  GTACACGGATTCCTCAGATATGAGTACTTCTAAGTTGATTTTCAGAAATATGAATATGAT 1800
Qy      1858  CTCAAATGGGAGTTTCAAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCT 1917
Db      1801  GTCAATGGGAGTTTCAAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCT 1860
Qy      1918  TTTGGAAAAGTGTATGAACGCAACAGCTTTATGAAATTTAGCAAAAACAGAGAGTCTCAATCCAG 1977
Db      1861  TTTGGAAAAGTGTATGAACGCAACAGCTTTATGAAATTTAGCAAAAACAGAGAGTCTCAATCCAG 1920
Qy      1978  GTTGGCGTCAAAATGCTGAAGAAAAGCAGAGCTCTGAAGAGAGGCACTCATGTCA 2037
Db      1921  GTTACCGTCAAAATGCTGAAGAAAAGCAGAGCTCTGAAGAGAGGCACTCATGTCA 1980
Qy      2038  GAACTCAAGATGATGACCGAGCTGGGAAGCCAGAGAAATTTGTAACCTGCTGGGGGGG 2097
Db      1981  GAACTCAAGATGATGACCGAGCTGGGAAGCCAGAGAAATTTGTAACCTGCTGGGGGGG 2040
Qy      2098  TGCACATCTCAGGACCAATTTACTGATTTTGAATTTGCTATGCTATGCTATCTTC 2157
Db      2041  TGCACATCTCAGGACCAATTTACTGATTTTGAATTTGCTATGCTATGCTATCTTC 2100
Qy      2158  AACTATCTAAGAGTAAAGAGAAAATTTCAAGAGCTTTGGACAGAGATTTTCAAGGAA 2217
Db      2101  AACTATCTAAGAGTAAAGAGAAAATTTCAAGAGCTTTGGACAGAGATTTTCAAGGAA 2160
Qy      2218  CACAATTTTCAAGTTTACCCCACTTTTCAATCATCAATTTCAATTTCCAGATGCTGTTCA 2277
Db      2161  CACAATTTTCAAGTTTACCCCACTTTTCAATCATCAATTTCAATTTCCAGATGCTGTTCA 2220
Qy      2278  AGAAGATTTCAATACACCGGAGCTCGGATCAATCTCAGGCTTTCATGGGATTCATTT 2337
Db      2221  AGAAGATTTCAATACACCGGAGCTCGGATCAATCTCAGGCTTTCATGGGATTCATTT 2280
Qy      2338  CACTCTGAAGATGAATTTGAATTTGAAAACCAAAAAGGCTGGAAGAGAGAGGACTTG 2397
Db      2281  CACTCTGAAGATGAATTTGAATTTGAAAACCAAAAAGGCTGGAAGAGAGAGGACTTG 2340
Qy      2398  AATGCTTTACATTTTGAAGATCTTTTCTGCTTGCATCAATAGTTGCCAAAGAGTGA 2457
Db      2341  AATGCTTTACATTTTGAAGATCTTTTCTGCTTGCATCAATAGTTGCCAAAGAGTGA 2400
Qy      2458  TTTCTGGAATTTAAGTCGTTGTTTCAAGAGACCTGGCGCCAGAGAACTGCTGTCACC 2517
Db      2401  TTTCTGGAATTTAAGTCGTTGTTTCAAGAGACCTGGCGCCAGAGAACTGCTGTCACC 2460
Qy      2518  CAGGGAAGTGGTGAAGATATGTGACTTTTGAATTTGGCTTCGAGATATCATGATGATTC 2577
Db      2461  CAGGGAAGTGGTGAAGATATGTGACTTTTGAATTTGGCTTCGAGATATCATGATGATTC 2520
Qy      2578  AACTATGTTTGCAGGGCAATGCCCGTCTGCTGTTAAATGATGAGCCGCCGAAAGCCTG 2637
Db      2521  AACTATGTTTGCAGGGCAATGCCCGTCTGCTGTTAAATGATGAGCCGCCGAAAGCCTG 2580
Qy      2638  TTTGAAGGCATCTACACCAATTAAGAGTGTGCTGCTCATATGGAATTTACTGTGGGA 2697
Db      2581  TTTGAAGGCATCTACACCAATTAAGAGTGTGCTGCTCATATGGAATTTACTGTGGGA 2640
Qy      2698  ATCTTCTCATTGTTGTAATCTTTACCTTGGCAATTCGGTTTGTATGCTAACTTTTACAAA 2757
Db      2641  ATCTTCTCATTGTTGTAATCTTTACCTTGGCAATTCGGTTTGTATGCTAACTTTTACAAA 2700
Qy      2758  CTGATTTCAAAATGATTTAAATGATGATGAGCAATTTTATGCTACAGAAAGAAATATACATT 2817
Db      2701  CTGATTTCAAAATGATTTAAATGATGATGAGCAATTTTATGCTACAGAAAGAAATATACATT 2760
Qy      2818  ATATGCAATCTCTGGGCTTTTGAATTTTGAAGCAAGCTCCCTGAAATTTGCCAGCAATTTTGA 2877
```

```
Db      2761  ATAATGCAATCTCTGGGCTTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAAT 2820
Qy      2878  TCGTTTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2937
Db      2821  TCGTTTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2880
Qy      2938  CGTGTTCGGAATGCTCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 2997
Db      2881  CGTGTTCGGAATGCTCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 2940
Qy      2998  TTGGGGCTACTCTCTCCGAGGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 3057
Db      2941  TTGGGGCTACTCTCTCCGAGGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 3000
Qy      3058  AGGACTTCATCTCTCCGAGGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 3117
Db      3001  AGGACTTCATCTCTCCGAGGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 3060
Qy      3118  CATCACTTAAAGAAAATCTTATTAATCACTGCTGCTTCCAGAGCTTTTCTTAGAGCG 3177
Db      3061  CATCACTTAAAGAAAATCTTATTAATCACTGCTGCTTCCAGAGCTTTTCTTAGAGCG 3120

RESULT 12
AX695523
LOCUS      2982 bp      DNA      linear      PAT 31-MAR-2003
DEFINITION      Sequence 1150 from Patent WO03008583.
ACCESSION      AX695523
VERSION      AX695523.1 GI:29418675
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1
AUTHORS      Morris,D.W. and Engelhard,E.K.
TITLE      Novel compositions and methods for cancer
JOURNAL      Patent: WO 03008583-A 1150 30-JAN-2003;
Sagres Discovery (US)
FEATURES
Source      1..2982
Location/Qualifiers
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
BASE COUNT      896 a 612 c 683 g 791 t
ORIGIN
Query, Match      85.2%; Score 2982; DB 6; Length 2982;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2982; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      58  ATGCCGGGTTGGCGCGGACGCGGCGGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 117
Db      1  ATGCCGGGTTGGCGCGGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 60
Qy      118  ATATTTGGGACTATTACAAATCAAGATCTGCTGCTGATCAAGTGTGCTTTTAAATCAATCAT 177
Db      61  ATATTTGGGACTATTACAAATCAAGATCTGCTGCTGATCAAGTGTGCTTTTAAATCAATCAT 120
Qy      178  AAGAACAAATGATTTCAATGAGTGGGAAAGTCATCATCATATCCCATGATCAGAAATCCCCG 237
Db      121  AAGAACAAATGATTTCAATGAGTGGGAAAGTCATCATCATATCCCATGATCAGAAATCCCCG 180
Qy      238  GAAGACCTCGGGTGTGCGGTTGAGACCCAGAGCTCAGGACAGTGTACGAAGTGGCGGT 297
Db      181  GAAGACCTCGGGTGTGCGGTTGAGACCCAGAGCTCAGGACAGTGTACGAAGTGGCGGT 240
Qy      298  GTGGAAGTGGATGTATCTGCTTCCATCACACTGCAAGTGGTGGTGCATGCCCGGGAAC 357
Db      241  GTGGAAGTGGATGTATCTGCTTCCATCACACTGCAAGTGGTGGTGCATGCCCGGGAAC 300
Qy      358  ATTTCTGCTCTCTGGGCTTTTGAAGCAAGCTCCCTGAAATTTGCCAGCAATTTTGAATTTA 417
```

Db 301 ATTTCTCTCTCGGGTCTTTAAGCAGCTCCCTGAAATTGCCAGCCACATTTTGATTTA 360  
Qy 418 CAAAAACAGAGGAGTTGTTTCCATGGTCAATTTTGAAAAATGACAGAAACCCCAAGCTGGAGAA 477  
Db 361 CAAAAACAGAGGAGTTGTTTCCATGGTCAATTTTGAAAAATGACAGAAACCCCAAGCTGGAGAA 420  
Qy 478 TACCTACTTTTATTACAGAGTGAAGCTACCAATTACCAAAATTTGTTTACAGTGAGTATA 537  
Db 421 TACCTACTTTTATTACAGAGTGAAGCTACCAATTACCAAAATTTGTTTACAGTGAGTATA 480  
Qy 538 AGAATACCTCTCTTACACATTTAGAAAGACCTTACTTTAGAAATGAAACCAAGGAC 597  
Db 481 AGAATACCTCTCTTACACATTTAGAAAGACCTTACTTTAGAAATGAAACCAAGGAC 540  
Qy 598 GCCCTGGTCTGCATATCTCAGAGCGTTTCCAGAGCCGATCGTGGAAATGGGTCTTTGCGAT 657  
Db 541 GCCCTGGTCTGCATATCTCAGAGCGTTTCCAGAGCCGATCGTGGAAATGGGTCTTTGCGAT 600  
Qy 658 TCACAGGGGAAAGCTGTAAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAGTG 717  
Db 601 TCACAGGGGAAAGCTGTAAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAGTG 660  
Qy 718 CTTCAATGAATTTTGGGACGACATTAAGTGTCTGTCAGAAATGAACTGGGACGGAA 777  
Db 661 CTTCAATGAATTTTGGGACGACATTAAGTGTCTGTCAGAAATGAACTGGGACGGAA 720  
Qy 778 TGCACAGGCTGTTCACAAATAGATCTAAATCAAACTCCTCAGACCAATTTGCCACAATTA 837  
Db 721 TGCACAGGCTGTTCACAAATAGATCTAAATCAAACTCCTCAGACCAATTTGCCACAATTA 780  
Qy 838 TTTCTTAAAGTAGGGGAAACCTTATGGATAAGTGCAAAAGCTGTTTCATGTGAACCATGGA 897  
Db 781 TTTCTTAAAGTAGGGGAAACCTTATGGATAAGTGCAAAAGCTGTTTCATGTGAACCATGGA 840  
Qy 898 TTGCGGCTCACCTGGGAAATTAGAAAACAAAGCACTCAGAGGGGCAACTACTTTGAGATG 957  
Db 841 TTGCGGCTCACCTGGGAAATTAGAAAACAAAGCACTCAGAGGGGCAACTACTTTGAGATG 900  
Qy 958 AGTACTATTCAACAAACAGAACTATGATACGGATTTCTGTTGCTTTTGTATCATCAGTG 1017  
Db 901 AGTACTATTCAACAAACAGAACTATGATACGGATTTCTGTTGCTTTTGTATCATCAGTG 960  
Qy 1018 GCAAGAAACGACACCGGATACACTGTTCTCTTCAAAGCATCCCAAGTCAATCAGCT 1077  
Db 961 GCAAGAAACGACACCGGATACACTGTTCTCTTCAAAGCATCCCAAGTCAATCAGCT 1020  
Qy 1078 TTGTTTACCATCGTAGGAAAGGATTTATAATGCTTACCAATTCAGAGTGAAGATTAAGAA 1137  
Db 1021 TTGTTTACCATCGTAGGAAAGGATTTATAATGCTTACCAATTCAGAGTGAAGATTAAGAA 1080  
Qy 1138 ATTGACCAATATGAAGATTTTCTTTTCTGTCAGGTTTAAAGCCTTACCACCAATCAGA 1197  
Db 1081 ATTGACCAATATGAAGATTTTCTTTTCTGTCAGGTTTAAAGCCTTACCACCAATCAGA 1140  
Qy 1198 TGTAGTGGACCTTCTCTCGAAAAATCAATTTCCCTTTGAGCAAAAGGCTCTTGATAACGGA 1257  
Db 1141 TGTAGTGGACCTTCTCTCGAAAAATCAATTTCCCTTTGAGCAAAAGGCTCTTGATAACGGA 1200  
Qy 1258 TACAGATATCCAAAGTTTGGCAATCATAGACACAGCCAGGAGGAATATATTTCCATGCA 1317  
Db 1201 TACAGATATCCAAAGTTTGGCAATCATAGACACAGCCAGGAGGAATATATTTCCATGCA 1260  
Qy 1318 GAAATATGATGATGCCCAATTTTACCAAAATGCTTCCAGCTGAATATAGAGGAACCTCAA 1377  
Db 1261 GAAATATGATGATGCCCAATTTTACCAAAATGCTTCCAGCTGAATATAGAGGAACCTCAA 1320  
Qy 1378 GTGCTCGCAGAGCATCGGCAAGTCAAGGCTCTGTTTCTCGAGTGGATACCAATTAACA 1437  
Db 1321 GTGCTCGCAGAGCATCGGCAAGTCAAGGCTCTGTTTCTCGAGTGGATACCAATTAACA 1380  
Qy 1438 TCTTGGACCTGGAAGAGTGTTCAGACAAAGTCTCCCAACTGCAAGAGATCAAGAA 1497  
Db 1381 TCTTGGACCTGGAAGAGTGTTCAGACAAAGTCTCCCAACTGCAAGAGATCAAGAA 1440

Qy 1498 GGAGTCTGGAATAGAAAGGCTAACAGAAAAGTGTTTGGA CAGTGGGTGTGAGCAGTACT 1557  
Db 1441 GGAGTCTGGAATAGAAAGGCTAACAGAAAAGTGTTTGGA CAGTGGGTGTGAGCAGTACT 1500  
Qy 1558 CTAACAATGAGTGAAGCCATAAAAAGGTTCTTGGTCAAGTGTGTCATACAAATCCCTT 1617  
Db 1501 CTAACAATGAGTGAAGCCATAAAAAGGTTCTTGGTCAAGTGTGTCATACAAATCCCTT 1560  
Qy 1618 GGCACATCTGTGAGACGATCCTTTTAAACTCTCCAGGCCCTTCCCTTTTATCCCAAGAC 1677  
Db 1561 GGCACATCTGTGAGACGATCCTTTTAAACTCTCCAGGCCCTTCCCTTTTATCCCAAGAC 1620  
Qy 1678 AACATCTCATTTCTATGCAACAAATTTGGTGTGTTGCTCTCTCTTCAATGTGCTTTTAAACCTG 1737  
Db 1621 AACATCTCATTTCTATGCAACAAATTTGGTGTGTTGCTCTCTCTTCAATGTGCTTTTAAACCTG 1680  
Qy 1738 CTAATTTGTCAAGTGAAGAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAG 1797  
Db 1681 CTAATTTGTCAAGTGAAGAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAG 1740  
Qy 1798 GTGACCGGCTCTCAGATAATGAGTACTTCTACGTTTGTATTTTCAAGAGAATATGATATGAT 1857  
Db 1741 GTGACCGGCTCTCAGATAATGAGTACTTCTACGTTTGTATTTTCAAGAGAATATGATATGAT 1800  
Qy 1858 CTCAAAATGGGAGTTTTCCAAAGAGAAAATTTTAGAGTTTGGGAAAGGTACTAGGATCAGGTGCT 1917  
Db 1801 CTCAAAATGGGAGTTTTCCAAAGAGAAAATTTTAGAGTTTGGGAAAGGTACTAGGATCAGGTGCT 1860  
Qy 1918 TTTGAAAGAGTATGAAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAG 1977  
Db 1861 TTTGAAAGAGTATGAAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAG 1920  
Qy 1978 GTTGGCTCABAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGACACTCATGTCA 2037  
Db 1921 GTTGGCTCABAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGACACTCATGTCA 1980  
Qy 2038 GAACTCAAGATGATACCCAGCTGGGAAGCCAGAGAATTTTGTGAACCTGCTGGGGGGG 2097  
Db 1981 GAACTCAAGATGATACCCAGCTGGGAAGCCAGAGAATTTTGTGAACCTGCTGGGGGGG 2040  
Qy 2098 TGCACATGTCAGGACCAATTTTATCTTGTATTTTGAATATCTGTTGCTATGGTGAATCTCTC 2157  
Db 2041 TGCACATGTCAGGACCAATTTTATCTTGTATTTTGAATATCTGTTGCTATGGTGAATCTCTC 2100  
Qy 2158 AACTATCTAAGAGTAAAGAGAAAATTTTCAAGAGCTTGGACAGAGATTTTTCAGGAA 2217  
Db 2101 AACTATCTAAGAGTAAAGAGAAAATTTTCAAGAGCTTGGACAGAGATTTTTCAGGAA 2160  
Qy 2218 CACAAATTCAGTTTTCACCCCACTTTTCCAAATCACTCCAAATTTCCAGCATGCCCTGGTTCA 2277  
Db 2161 CACAAATTCAGTTTTCACCCCACTTTTCCAAATCACTCCAAATTTCCAGCATGCCCTGGTTCA 2220  
Qy 2278 AGAAGATTCAGATACACCCGACCTGGATCAAAATCTCAGGCTTCAAGGCTTCAAGGAAATCATTT 2337  
Db 2221 AGAAGATTCAGATACACCCGACCTGGATCAAAATCTCAGGCTTCAAGGCTTCAAGGAAATCATTT 2280  
Qy 2338 CACTCTGAAGATGAAATTTGAATATGAAAACCAAAAAGGCTGGAGAAAGGAGGAGCTTG 2397  
Db 2281 CACTCTGAAGATGAAATTTGAATATGAAAACCAAAAAGGCTGGAGAAAGGAGGAGCTTG 2340  
Qy 2398 AATGTCTTACATTTTGAAGATCTTTCTGCTTTGCAATATCAAGTTCGCAAGGAAATGGAA 2457  
Db 2341 AATGTCTTACATTTTGAAGATCTTTCTGCTTTGCAATATCAAGTTCGCAAGGAAATGGAA 2400  
Qy 2458 TTTCTGGAATTTAAGTGTGTTTCAAGAGCTTGGCCGCCAGGAAACGCTGTTGTCAAC 2517  
Db 2401 TTTCTGGAATTTAAGTGTGTTTCAAGAGCTTGGCCGCCAGGAAACGCTGTTGTCAAC 2460  
Qy 2518 CACGGGAAAGTGGTGAAGATATGTGATCTTTGATTTGGCTTCGAGATATCATGATGATTCC 2577  
Db 2461 CACGGGAAAGTGGTGAAGATATGTGATCTTTGATTTGGCTTCGAGATATCATGATGATTCC 2520



1361 AGTTCCAAAATGTTTCAACCTGAAATATAAGAAAGAAACCTCAAGTCTAGCAAAATGCCT 1420  
1394 CGGCAAGTCAGGCGTCTGTTTCTCGATGGATACCATACCATCTTTGACCTTGGGAAGA 1453  
1421 CAGCCAGCCAGGCGTCTGTTTCTGATGGCTACCCGCTACCTCTTTGACCTTGGGAAGA 1480  
1454 AGTGTTCAGACAAAGTCTCCCAAATGACACAGAAAGATCACAGAAAGGAGTCTGGAAATAGAA 1513  
1481 AGTGTTCGGACAAATCTCCCAAATGACACAGAAAGATCCCAAGAAAGGAGTCTGGAAATAGAA 1540  
1514 AGGCTAACAGAAAGTCTTTGGGACAGTGGGTGTCAGAGCTACTCTTAAACATGAGTGAAG 1573  
1541 AGGCTAACAGAAAGTCTTTGGGACAGTGGGTGTCAGAGCTACTCTTAAATATGAGTGAAG 1600  
1574 CCATATAAGGGTCTTGGTCAAGTGTGTCATACAAATTCCTTGGGCACATCTTGTGAGA 1633  
1601 CCGGGAAGGGCTTCTGGTCAAAATGCTGTGCGTACAAATTCATTTGGGCACGCTTTTGGGAAA 1660  
1634 CGATCTCTTTTAAACTCTCCAGGGCCCTTCCCTTTTCATCCAGAGACAAATCTCAATCTTATG 1693  
1661 CCATCTTTTAAACTCACAGGGCCCTTCCCTTTTCATCCAGAGACAAATCTCTCTTCTATG 1720  
1694 CAACAATTTGGTGTCTCTCTTCTATGTCGTTTAAACCTGTAAATTTGTCTCAAAAT 1753  
1721 CGACCAATGGGCTCTGTCTCCCTTCAATGTTGTTCTCATTTGTTGATCTGTCCACAAAT 1780  
1754 ACBAAGCAATTTAGGTATGAAGCAGCTACAGATGGTACAGGTGACCGGCTCTCTCAG 1813  
1781 ACBAAGCAATTTAGGTATGAGAGTACAGATGGTACAGGTGACCGGCTCTCTCAG 1840  
1814 ATAAATGAGTACTTCTACGTTGATTTACAGAAATATGAAATATGATCTCAATTTGGAGTTTC 1873  
1841 ATAAACGAGTACTTCTACGTTGATTTACAGAGTATGAAATATGATCTCAATTTGGAGTTTC 1900  
1874 CAAAGAGAAATTTAGAGTTTGGAGGTTACTAGATCAGTGTCTTTTGGAAAGTGTATGA 1933  
1901 CGAGAGAGAACTTAGAGTTTGGAGGTTCTCTGGGCTCTGGGCTTTTGGAGGTTGTATGA 1960  
1934 ACCCAACAGCTTTAGAAATTTAGCAAAACAGAGTCTCAATCCAGGTTTCCGCTCAAAATGC 1993  
1961 ACCCAACAGCTTTAGCAATTTAGTAAACAGGAGTCTCAATTTAGGTTTGGAGGTTGTATGA 2020  
1994 TGAAGAGAAAGCAGACAGCTCTGAAAGAGAGGACCTCATGTACAGAACTCAAGATGATGA 2053  
2021 TAAAGAGAAAGCTGACAGCTGTGAAAGAGAGCTCTCATGTGCGAGCTCAAAATGATGA 2080  
2054 CCCAGCTGGAGAACCCAGAGAAATTTGTGAACTGTGCTGGGGGCGTGACACTGTGAGGAC 2113  
2081 CCCACCTGGGACACCATGACAAATCGTGAATCTGCTGGGGGCGATGCACACTGTGAGGCG 2140  
2114 CAATTTACTTTGATTTTGAATCTGTTGCTATGTTGATCTTCTCAATCTATTAAGAGTGA 2173  
2141 CAGTGTACTTTGATTTTGAATTTGTTGCTATGTTGACCTCTCACTACTCTAAGAGTGA 2200  
2174 AAAGAGAGAAATTTACAGGACTTTGACAGAGATTTTCAAGGAGACAAATTTCAAGTTTTC 2233  
2201 AAAGAGAGAAATTTACAGGACATGGAAGAGATTTTAAAGAGACAAATTTCAAGTTTTC 2260  
2234 ACCCCACTTTCCAAATCATCTCAAAATTTCCAGCATGCTGTTTCAAGAGAGAGTTTCAAGATC 2293  
2261 ACCCTACTTTCCAGGACATCTCAAAATTTCCAGCATGCTGTTTCAAGAGAGAGTTTCAAGTTC 2320  
2294 ACCCGACTCGGATCAAAATCTCAGGCTCTCATGGAATTTCAATTTCACTCTGAAGATGAA 2353  
2321 ACCCGCTCTGGATCAGCTCTCAGGCTTCAATGGGAAATTTCAATTTCACTCTGAAGATGAA 2380  
2354 TTGAATATGAAGAACCAAGAGGCTG-----GAAGAGAGGAGGACTTGAATGCTGCTTA 2407  
2381 TTGAATATGAAGAACCAAGAGGCTGCGAGAGAGAGGAGAGATTTGAACTGCTGTA 2440  
2408 CATTTGAAGATCTTCTTTGCTTTGCTATATCAAGTTTGCACAAAGGAATTTCTTGAAT 2467

2441 CGTTTGAGACCTCTCTTTGCTTTGCGTACCAAGTGCCCAAGGCAATGGAATTTCTGTGAGT 2500  
2468 TTAAGTCGTGTGTTTACAGAGACTCTGGCCGCCAGGAACGTGCTTTGTCAACCCACGGGAAG 2527  
2501 TCAAGTCGTGTGTCACAGAGACTCTGGCAGCAGGAATGTGTTGGTCAACCCACGGGAAG 2560  
2528 TGTGTAAGATATGTAATTTGGATTTGGCTCGAGATATCATGAGTGAATCCAACTATGTTG 2587  
2561 TGTGTAAGATCTGTGACTTTGGACTCTGGCCCGAGACATCTCTGAGCGACTCCAGCTACGTCG 2620  
2588 TCAGGGGCAATGCCGTCTGCTGTAATAATGGATGGCCCCCGGAAAGCCTGTTTCAAGGCA 2647  
2621 TCAGGGGCAACGCAACGCTGCCGTGAAATGGATGGCACCCGAGAGCTTATTTGAAGGGA 2680  
2648 TCTACACCAATTAAGAGTGTCTGTGTCATATGGAATTAATTAATTTGAGGAAATCTTCTCAC 2707  
2681 TCTACACCAATCAAGAGTGACGTCTGTGTCCTACGGCATCTCTCTGAGGAGATATTTTTCAC 2740  
2708 TTGGTGTGAATCTTACCTTGGCATTTCCGGTTGATGCTAACTTCTCAAACTGATTTCAA 2767  
2741 TGGGTGTGAACCTTACCTTGGCATTTCTGTGCGACGCTAACTTCTATATAAACTGATTCAGA 2800  
2768 ATGGATTTAAAAATGGATCAGCCATTTTATGCTACAGAAAGAAATATATATATATATGAAT 2827  
2801 GTGGATTTAAAAATGGAGAGCCATTTCTATGCCACAGAAAGGATATATCTTTGTAATGAAT 2860  
2828 CTTGCTGGGCTTTGACTCAAGGAAACGGCCATCTCTTCCCTAAATTTTGACTTTGTTTTAG 2887  
2861 CTTGCTGGGCTTTGACTCAAGGAAACGGCCATCTCTTCCCAACCTTGACTTTCATTTTAG 2920  
2888 GATGTCAAGTGGCAGATGCAAGAAAGCGATGATACAGAAATGTTGGATTTGGCGGTGTTTCGG 2947  
2921 GATGTCAAGTGGCAGAGGCAAGAAAGCGATGATACAGAAATGTTGGGTGGCAACGTCGCCAG 2980  
2948 AATGTCTCACACTTACCAACAGGCGACTTTTACAGAGAGATGATTTGGGGGTAC 3007  
2981 AACATCTCATCTTACCAAAACAGGCGCTCTTACAGAGAGGCGGCTCAGAGCGCGC 3040  
3008 TCTCTCCGAGGCTCAGGTCAAGATTTCTAGAGGAAATTTAGTTTAAAGGACTTTCAT 3067  
3041 CATGCGCAACGCGCCAGGT--GAAGATTTACAGAGAAAGATTTAGCGAGGAGGCTTTGGA 3099  
3068 CCCTCCAACCTTATCCCTAAACAGGCTGTAGATTAACAAAACAGATTAATTTCACTCAAAA 3127  
3100 CCCCAGCA---CCCTAGCAGGCTGTAGACCGCAGAGCAAGATTTAGCCTCGCCTCT--G 3153  
3128 AGAAATCTATATTAATCACTGCTGCTTCAACAGCTTTTCTAGAGAGCGCTGCGGTTTA 3187  
3154 AGGAAGCGCCCTACAGGCGCTTGTCTGCTGGAGCTTTTCTCTAGATGCTGTCTGCCATTA 3213  
3188 CTCTTTGTTTCAAGGGAGCTTTTGTAAATCAAAATCATCTGTCAAGGCGAGGAGGAGC 3247  
3214 CTC-----CAAAGTGACTTTCTATAAAATCAAACTCTCTCTGCAAGGCGGAGAGGCC 3266  
3248 TGATAATGAATTTTATTTGGAGCAATGATCTGCAATCCAAAGGCTTTCTCAGGCGCGCTTTGAG 3307  
3267 AATAATGAGACTTTGTTGGTGGCCCGCTTACCCTTGGGGGCGCTTTTCCAGGCGCCCCAGCG 3326  
3308 TGAATTTG-----TGTAACCTGAAGTACAGTATATTTCTTTGTAATAATACATAAAACAAA-- 3357  
3327 TTGAGGGGAAAGCCATGTATCTGAAATATAGTATATTTCTTTGTAATAATACGTGAACCAACC 3386  
3358 -----AGCAATTTGCTTAAGGAGAGCTTAATGATTTTTT--AAGCTATGTTTTTAAAT 3410  
3387 AAACCCGTTTTTTTGTCTAAGGAAAGCTTAAATATGATTTTTTAAAAATCTATGTTTTTAAAT 3446  
3411 AATATGTAAATTTTTTCAAGCTTATTTAGTGAATATTTTATGGGTGGGAATAAAAATTTCTAC 3470  
3447 ACTATGTAATTTTTTCACTATTTTAGTGAATATTTTATGATGAAATATAAACTTTCTAC 3506  
3471 TACAGAAAAA 3485  
3507 TGTAGAAAAA 3521

RESULT 14  
LOCUS  
MMFLT3  
DEFINITION  
Mouse Flt3 mRNA for tyrosine kinase receptor of the PDGF.  
ACCESSION  
X59398  
VERSION  
X59398.1 GI:50978  
KEYWORDS  
Flt3 gene; tyrosine kinase receptor.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
REFERENCE  
1  
AUTHORS  
Rosnet, O., Marchetto, S., delapeyriere, O. and Birnbaum, D.  
TITLE  
Murine Flt3, a gene encoding a novel tyrosine kinase receptor of  
the PDGFR/csf1r family  
JOURNAL  
Oncogene 6 (9), 1641-1650 (1991)  
MEDLINE  
92019834  
PUBMED  
1656368  
REFERENCE  
2 (bases 1 to 1321)  
AUTHORS  
Rosnet, O.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (07-MAY-1991) O. Rosnet, Cancerologie et therapeutique  
exp., Inst National de la Sante, 27 Boulevard Lei Rouse, 13009  
Marseille, France  
FEATURES  
Location/Qualifiers  
1..3521  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
1..3521  
/genes="Flt3"  
1..3521  
/genes="Flt3"  
/evidence="experimental"  
82..3084  
/genes="Flt3"  
/codon\_start=1  
/product="Flt3 protein"  
/protein\_id="CAA42041.1"  
/db\_xref="GI:50979"  
/db\_xref="MGI:95559"  
/translation="MRLAQSDRRLLLVLSMILEVTWVNDLPVIRKVLISHENN  
GSSAGKPSYVRVSGPEDLOCTPRROSGTVVEAATVEASGSLTLOVOLATPCDL  
SCLWFKHSLCQPHFDLQNGIUSMALINVTETQAGEYLLHIQSEANVYVLTAV  
VDTQLYVLRFRYFRMENQDALLCISEGPEPTVEWVLCSSHRESCKEGPAVRKE  
EVLHFLGTDIRCCARNALGRESTKLTFTDLNAPQSTLPQLFLKVGPELWIRKAI  
HYNGFGLTWELDKALEGYSFEMSTYSTRMTIRILLAFVSSVRNDTGYTCS  
KHPQSALVILEKGFINATSOEYEDIPYBKFCPSVRKAYPRIRCTWIFSOA  
CEQGLDGYSTSKFCDHKNKPEYIFYAENDDAOTFKMTFLNIRKPOVLANASAO  
ASCSGDPVLPSTWKKSDKSPNTEETEEIPEGWKKANRKYFGVWSSTLNSEAG  
KGLLVKCAYNMGSCTETIPLNSPGFFFIQNTISFYATIGLCUPFIIVLVILCHK  
YKQRYESQLQMIQVTPGLDNEYFYVDFRDYEDLKWEPFRENLEFVGLVSGAFGR  
VNNATYAGISGYSIQVAKMLKEKADSCKEALMSELKMTHLGHNDNI VNLGAC  
TUSGVPVLIFFCCYCDLLNLYRSKREKFTWTEIFKHNFSFYPTFOAHNSMPG  
SREVOLHPLDOLGPNHNSIHSDEIEYENOKRLAEEREEDNLVLTPEDLICFAYOV  
AKGMELFEKCVHRLDARNVLTGHVKVKICDFGLARDILSDSYVVRGNRLPVK  
WNAPSLPEGITIKSDVWSYGILLWEIFSLGNVPYPIGPVDANFYKLIQSGFKMEQP  
FYATGTYFVMSQWAFDSRKRPFPNLTSLFLGCLQALAEAEAMYQNMGNVPEHPSIY  
QNRRLSREAGSEPPSQAVKIHRS"  
BASE COUNT 943 a 854 c 907 g 817 t  
ORIGIN  
Query Match 65.0%; Score 2274.6; DB 10; Length 3521;  
Best Local Similarity 80.3%; Pred. No. 0;  
Matches 2805; Conservative 0; Mismatches 649; Indels 41; Gaps 10;  
Qy 17 GGCTGGCGCGCGCTCGGGGACCCCGGCTCCCGAGGCCATGCCGGCGTGGC---CC 73  
Db 42 GGCTGGCGCGCGCTCGCTACCGCGCTCCCGAGGCCATGCCGGCGTGGCGGCGAGC 100

Qy 74 GCGACGGGGCACCGTCCCGCTGCTCGTTGTTTTCTGCAATGATATTTGGGACTATTA 133  
Db 101 GCAGCGACCGGGCGGCTGCTGCTGTTGTTTTGTGAGTAATGATTTTGGAGACCGTTA 160  
Qy 134 CAAATCAAGATCTGCTGTGATCAAGTGTGTTTTTAATCAATCAATAAGAAACAATGATTCAT 193  
Db 161 CAAACCAAGACCTGCTGTGATCAAGTGTGTTTTTAATCAATCAATAAGAAACAATGATTCAT 220  
Qy 194 CAGTGGGGAAGTCAATCATCATATCCCATCGTATCAGAATCCCGGGAAGACCTCGGGTGTG 253  
Db 221 CAGCGGGAAGGCCATCATCGTACCGAATGGTGCAGGATCCCGAGAGACCTCCAGTGTA 280  
Qy 254 CGTTGAGACCCAGAGCTCAGGACAGTGTACAGAGTGGCGCTGTGGAAGTGGATGTAT 313  
Db 281 CCCGAGGCGCCAGAGTGAAGGACCGGTATATGAAGCGGCCACCGTGGAGTGGCCAGT 340  
Qy 314 CTGCTTCCATCACACTGCAAGTCTGCTGATCGATCCCCAGGGAACATTTCTGCTCTCGGG 373  
Db 341 CTGGGTCCATCACCTTGCAGTGCAGTCCGCCACCCAGGGGACCTTTCTGCTCTCGGG 400  
Qy 374 TCTTTAAGCACAGCTCCCTGAATTTGCCAGCCACATTTTGATTTTACAAAAACAGAGAGTTG 433  
Db 401 TCTTTAAGCACAGCTCCCTGGGCTGCCAGCGCACTTTGATTTTACAAAAACAGAGAGTTCG 460  
Qy 434 TTTCCATGTCTATTTTGAATAACAGAAACCCAGCTGGAGAAATACCTACTTTTATTC 493  
Db 461 TTTCCATGTCCATCTTGAACGTGCAGAGACCCAGGAGGAGAAATACCTACTTCCATATTC 520  
Qy 494 AGAGTGAAGCTACCAATTACAGAAATTTTACAGTGTATTAAGAAATACCTGCTTT 553  
Db 521 AGAGCAAGCCCGCAACTACAGTACTGTTCACAGTGAATTAAGAGAGATACACAGCTGT 580  
Qy 554 ACACATTAAGAAAGACCTTACTTTAGAAAAATGAAAAACAGGACGCCCTGCTCTGCATAT 613  
Db 581 ATGTGCTAAGGAGACCTTACTTTAGGAGATGAAAAACAGGATGCACTGCTCTGCATCT 640  
Qy 614 CTGAGAGCGTTCCAGAGCGCGATCGTGGAAATGGGTGCTTTGCGATTCACAGGGGAAAGCT 673  
Db 641 CCGAGGGTGTTCGGAGGCCCACTGTGGAGTGGGTGCTCTGCAGCTCCCAAGGAAAGCT 700  
Qy 674 GTAAAGAAAGAGTCCAGCTGTTGTTAAAGAGGAGAAAAAGTGTTCATGAATTTATTTG 733  
Db 701 GTAAAGAAAGAGGCCCTGCTGTTGTTCAGAAAGAGGAAAAAGGTACTTTCATGAGTTGTCG 760  
Qy 734 GGACGGACATAAGTGTCTGTCAGAAATGAATGGGCGAGGGAATGCACAGAGCTGTTC 793  
Db 761 GAAACAGACATCAGATGCTGTGTAGAAATGCATGGGCGCGGAAATGCACCAAGCTGTTC 820  
Qy 794 CAATAGATCTAAATCAAACTCCCTCAGACCCACATTTGCCAAATTTATTTCTTAAAGTAGGGG 853  
Db 821 CCATAGATCTAAACCAAGGCTCTCTCAGAGACACACTGCCCCAGTTATTTCTGAAAGTGGGG 880  
Qy 854 AACCTTATGGAATAAGGTGCAAGCTGTTTCAATGAAACCATGGATTCGGGCTCACCTGGG 913  
Db 881 AACCTTGTGGATCAGGTGTAAGGCCATCCATGTGAAACCATGGATTCGGGCTCACCTGGG 940  
Qy 914 AATTAGAAACAAAGCACTCGAGGAGGCAACTACTTTGAGATGATGATCTATTCAACAA 973  
Db 941 ACTGGAAGACAAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGATGATCTACTTCCACAA 1000  
Qy 974 ACAGAACTATGATACGGATTCTGTTGCTTTTGTATCATCATGTGGCAAGAAACGACACCG 1033  
Db 1001 ACAGGACCATGATTCGGATTCTTTGGCCTTTGTGCTTCCTGGGAAAGAACGACACCG 1060  
Qy 1034 GATATCACTTGTTCCTTCCAAAGCATCCAGTCAATCAGCTTTGGTTGGTACCATCGTAG 1093  
Db 1061 GATATTAACCTGCTCTTCTCCTCAAAGCACCCAGCCAGTCCAGCTTTGGTGGACCATCCTAG 1120  
Qy 1094 GAAAGGATTTATAATGCTACCAATTCAGTGAAGATTTAATTAATGACCAATATGAAG 1153  
Db 1121 AAAAGGGGTTTATAAACGCTACAGCTCGCAAGAGAGATGAATGAATGACCCGCTACGAAA 1180  
Qy 1154 AGTTTGTGTTTCTGTGAGGTTTAAAGCCTACCCCAAAATCAGATGATGATGAGACCTTCT 1213



1181 AGTTCTGCTTCTCAGTCAGGTTTAAAGCGTACCCAGAAATCCGATGCACTGATCTTCT 1240  
1214 CTCGAAAAATCATTTCTCTTGAGCAAAAGGCTCTTGATACGATACAGCATATCOAGT 1273  
1241 CTCAGGCTCATTTCTCTTGAGCAAGAGGCTTGAGGATGAGTACAGCATATCTAAT 1300  
1274 TTTGCAATCATAGCACAGCCAGGAGAAATATATTTCCATGAGAAAAATGATGCCCC 1333  
1301 TTTGCGATCATAGAAACAGCCAGGAGTACATATTTCTATGAGAAAAATGATGCCCC 1360  
1334 AATTACCAAAATGTCAGCTGAATATAGAGAGAAACCTCAAGTCTCGAGAGCAT 1393  
1361 AGTTCAACCAAAATGTTCACTGCTGATATAGAGAGAAACCTCAAGTCTAGCAATGCTCT 1420  
1394 CGCAAGTCAGGCTGCTGCTCTCGGATGATACCAATTTACCATTTTGACCTTGAAGA 1453  
1421 CAGCCAGCCAGGCTGCTGCTCTCTGATGGCTACCCGCTACCCCTTTTGACCTTGAAGA 1480  
1454 AGTGTTCAGACAAATCTCCCAATCGCAAGAGAGATCACAGAGGAGTCTGGAATAGAA 1513  
1481 AGTGTTCGGACAAATCTCCCAATTTGCAAGGAGAAATCCAGAGGAGTTTGGAAATAAA 1540  
1514 AGCTTAACAGAAAAATGTTGGACAGTGGGTGTCGAGCATCTCTAAACATGATGAG 1573  
1541 AGCTTAACAGAAAAATGTTGGCCAGTGGGTGTCGAGCATCTCTAAATATGATGAGG 1600  
1574 CCATAAAGGTTCTGCTCAAGTGTGTCATCAATTTCCCTTGGCACATCTTTGTAGA 1633  
1601 CCGGAAAGGCTCTGCTGCTCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1660  
1634 CGATCTCTTTAAATCTCTCAGGCCCCCTTCCCTTTTCATCCCAAGACAAATCTCTATG 1693  
1661 CCATCTTTTAAATCTCAGGCCCCCTTCCCTTTTCATCCCAAGACAAATCTCTATG 1720  
1694 CAACAAATGTTGTTGCT 1753  
1721 CGACCAATGCGCTCTGCT 1780  
1754 ACAAAGCAATTTAGTATGAAGCCAGCTACAGATGTTACAGGTGACCGGCTCTCTCAG 1813  
1781 ACAAAGCAATTTAGTATGAAGCCAGCTACAGATGTTACAGGTGACCGGCTCTCTCAG 1840  
1814 ATAAAGCAATTTAGTATGAAGCCAGCTACAGATGTTACAGGTGACCGGCTCTCTCAG 1873  
1841 ATAAAGCAATTTAGTATGAAGCCAGCTACAGATGTTACAGGTGACCGGCTCTCTCAG 1900  
1874 CAAAGAAAAATTTAGATTTGGAAAGGTACTAGATCAGGTGCTTTTGGAAAAATGATGA 1933  
1901 CGAGAGAGAACTTAGAGTTTGGAAAGGCTCTGCGGCTCTGCGGCTCTGCGGAGGTGATGA 1960  
1934 ACGCAACAGCTTATGNAATTTAGCAAAAACAGAGTCTCAATCCAGGTGCGGCTCAAAATGC 1993  
1961 ACGCCAAGCTTATGNAATTTAGTATGAAGCCAGGAGTCTCAATTTAGGAGGTGAGATGC 2020  
1994 TGAAGAAAAACAGACAGCTCTGAAAGAGAGGCACTCATGTACAGAACTCAAGATGATGA 2053  
2021 TAAAGAGAAAGCTGACAGCTGTGAAAGAGAGCTCTCATGTGCGAGCTCAAAATGATGA 2080  
2054 CCCAGCTGGAAGCCACGAGAAATATTTGTGAACCTGCTGGGGCGGTGACACTGTCAGGAC 2113  
2081 CCCACCTGGGACCAATGACAACTCTGTAATCTGCTGGGGGCAATGACACTGTCAGGCG 2140  
2114 CAATTTACTTTGATTTTGAATCTGTTGCTATGTTGATCTCTCAACTATCTAAGAGTA 2173  
2141 CAGTGTACTTTGATTTTGAATATTTGTTGCTATGTTGATCTCTCAACTATCTAAGAGTA 2200  
2174 AAAGAGAAAAATTTACAGGACTTGGACAGAGATTTTCAAGGAACACAAATTTCAAGTTTTT 2233  
2201 AAAGAGAGATTTTACAGGACATGGAACAGAGATTTTAAAGGAACATAAATTTCAAGTTTTT 2260  
2234 ACCCACTTTTCAATCACTCCAAATTTCCAGCATGCTGTTTCAAGAGAGATTTCAAGATAC 2293

2261 ACCCTACTTTTCAGGGCACATTTCAAAATTCAGCATGCTGTTTCAGAGAAATTCAGTTAC 2320  
2294 ACCCGACTCGGATCAAAATCTCAGGGCTTTCTAGGGAATTCATTTCACTCTGAAAGTAA 2353  
2321 ACCCGCTTTGGATCAGCTCAGGCTTCAATGGGAATTCATTTCTCTGAAAGTGA 2380  
2354 TTGAATATGAAAAACCAAAAAAGGCTG-----GAAAGAGAGGAGGACTTGAATGTGCTTA 2407  
2381 TTGAATATGAAAAACCAAGAGAGGCTGCGAGAAAGAGAGGAGGAATTTGAAACGTGCTGA 2440  
2408 CATTTCAAGATCTTTCTTTGCTTTGCTATCAAGTGTGCAAGGAAATGGAATTTCTGGAAT 2467  
2441 CGTTTCAAGACCTTCTTTGCTTTGCTATCAAGTGTGCAAGGAAATGGAATTTCTGAGT 2500  
2468 TTAAGTCTGTTGTTCACAGAGACTGCGCGCCAGAGAACTGCTTGTTCACCCAGCGGAAAG 2527  
2501 TCAAGTCTGTTGTTCACAGAGACTGCGCGCCAGAGAAATGTTGTTGTTGTTGTTGTTGTTG 2560  
2528 TGGTGAAGATATGTGACTTTTGGATTTGGCTGAGATATCATAGATGATTTCCAACTATGTTG 2587  
2561 TGGTGAAGATATGTGACTTTTGGATTTGGCTGAGATATCTCTGAGGCACTCCAGCTACGTCG 2620  
2588 TCAGGGGCAATGCGCTCTGCTGTAATGATGCGCCCGGAGAGGCTGTTTGAAGCA 2647  
2621 TCAGGGGCAATGCGCTCTGCTGTAATGATGCGCCCGGAGAGGCTTATTTGAAGGGA 2680  
2648 TCTACACCAATTAAGAGTGTCTGCTGTAATGAGATATTTACTGTGGGAAATCTTCTCAC 2707  
2681 TCTACACCAATTAAGAGTGTCTGCTGTAATGAGATATTTACTGTGGGAAATCTTCTCAC 2740  
2708 TGGTGTGAATATGCTTACCCTGCGATTTCCGGTTGATGCTAACTCTTCAAACTGATTTCAA 2767  
2741 TGGTGTGAATATGCTTACCCTGCGATTTCCGGTTGATGCTAACTCTTCAAACTGATTTCAA 2800  
2768 ATGGAATTTAAATGGAATGAGCTTATGCTACAGAGAAATATATATATATATGCAAT 2827  
2801 GTGGATTTAAATGGAATGAGCTTATGCTACAGAGAAATATATATATATATGCAAT 2860  
2828 CTGCTGGGCTTTTGAATCAAGGAAACGGCACTCTTCCCTAAATTTGACTTTGCTTTTGTAG 2887  
2861 CTGCTGGGCTTTTGAATCAAGGAAACGGCACTCTTCCCTAAATTTGACTTTGCTTTTGTAG 2920  
2888 GATGTCAGCTGCGAGATGCAAGAAAGCGATGATCAGAAATGTTGATGCGGCTGTTTCGG 2947  
2921 GATGTCAGCTGCGAGAGGCAAGAAAGCGATGATCAGAAATGTTGATGCGGCTGTTTCGG 2980  
2948 AATGTCCTCACACTTACCAAAACAGGGGACTTTTTCAGCAGAGAGATGATTTGGGGCTAC 3007  
2981 AACATCCATCTTACCAAAACAGGGGACTTTTTCAGCAGAGAGATGATTTGGGGCTAC 3040  
3008 TCTCTCGGAGGCTCAGGTCGAAGATTTGTAAGGAAATTTAGTTTAAAGGACTTTCAT 3067  
3041 CATCGGCAAGGCTCAGGTCGAAGATTTGTAAGGAAATTTAGTTTAAAGGACTTTCAT 3099  
3068 CCCTCCACTTATCCCTAAACAGGCTGTAGATTTACCAAAACAGGATTTAAATTTTCACTTAA 3127  
3100 CCCCGCA-----CCCTAGCAGGCTGTAGACCGCAGAGGCAAGATTTAGCTCGCTCT--G 3153  
3128 AGAAATCTATATATCAACTGCTGCTTCAACAGCTTTTCTCTAGAGCGCTCTGCGTTTA 3187  
3154 AGGAAGCGGCTTACAGGCGGCTTTGCTGCTGAGGCTTTTCTCTAGATGCTGCTGCTGCTTA 3213  
3188 CTCTGTTTTTCAAGGAGCTTTTGTAAATCAATCATCTGTCACAGGAGGAGGAGC 3247  
3214 CTC-----CAAGTGAATTTTATATAATCAAACTCTCTCGCAAGGCGGAGAGCC 3266  
3248 TGATAATGAATTTTATTTGAGCAATTTGATCTGCAATTTCAAGGCTTCTCAGGCGGCTTGA 3307  
3267 AATAATGAGACTTTGTTGTTGAGCGCCCTTACCTGCGGGGCTTTTCCAGGCGCCCGAGC 3326  
3308 TGAATTTG-----TGTACCTGAAGTACAGTATATTTCTTTGTAATAATCATATAAACA 3357  
3327 TTGAGGGGAAAGCCATGATCTGAAATATAGTATATTTCTTTGTAATAATACGTGAAACAACC 3386





1652 QY CAGGCCCTTCCCTTTCATCCAAAGACAA CATCTCAATTCATGCAACAAATGGTGTGTC 1711  
1658 Db CAGGCCCTTCCCTTTCATCCAAAGACAA CATCTCCCTTCTATGCGACCAATGGGCTCTGTC 1687  
1712 QY TCCTCTTCATTCCTGTTTAAACCTGCTAATTTGTCACAAAGTACAAAGCAATTTAGGT 1771  
1688 Db TCCTCTTCATTCCTGTTTAAACCTGCTAATTTGTCACAAAGTACAAAGCAATTTAGGT 1747  
1772 QY ATGAAGCCAGCTACAGATGGTACAGGTGACCGGCTCCCTCAGATAATGAGTACTCTACG 1831  
1748 Db ACAGAGTACAGTGCAGATGATCCAGGTGACCTGCGCCCTCGATACAGTACTCTACG 1807  
1832 QY TTGATTTACAGAAATATGAATATGATCTCAAAATGGAGTTTCCAAAGAAAATTTAGAGT 1891  
1808 Db TTGATTTACAGAAATATGAATATGATCTCAAAATGGAGTTTCCAAAGAAAATTTAGAGT 1867  
1892 QY TTGGGAAGTACTAGATCAGGTGCTTTTGGAAAAGTGTGATGACGCAACAGCTTATGGAA 1951  
1868 Db TTGGGAAGTCTCGGGTCTGGGCTTTTCGGGAGGTGATGAACGCGCCGCTATGGCA 1927  
1952 QY TTAGCAAAAACAGGAGTCTCAATCCAGGTTGCCGTCAAAATGCTGAAAGAAAAGCAGACA 2011  
1928 Db TTAGTAAACCGGAGTCTCAATTCAGGTGGGCTGAAGTGTCTAAAGAGAAAGCTGACA 1987  
2012 QY GCTCTGAAAGAGGCACTCATGTGCAAACTCAAGATGATGACCCAGCTGGGAAGCCAG 2071  
1988 Db GCTGTGAAAAGAGGCTCTCATGTGCGAGCTCAAAATGATGACCCAGCTGGGACACCATG 2047  
2072 QY AGAATATGTGAACCTGCTGGGGGCTGCACACTGTCCAGGACCAATTTACTGATTTTG 2131  
2048 Db ACAACATCTGTAATCTGCTGGGGGCAATGCACACTGTCCAGGGCCAGTGTACTGATTTTG 2107  
2132 QY AATACCTGTTGCTATGTGTATCTCTCAACTATCTAAGAAAGTAAAGAGAAAATTTTACA 2191  
2108 Db AATATTTGTTGCTATGTGTATCTCTCAACTATCTAAGAAAGTAAAGAGAAAATTTTACA 2167  
2192 QY GGAATTTGACAGAGATTTTCAAGAAACAAATTTCAATTTTACCCCACTTTCCAAATCAC 2251  
2168 Db GGAATTTGACAGAGATTTTCAAGAAACAAATTTTCAATTTTACCCCACTTTCCAGGCAC 2227  
2252 QY ATCCAAATTTCCAGCATGCTGTTCAAGAGAGTTTCAGATACACCCGGCTCCGATCMAA 2311  
2228 Db ATTCAAATTTCCAGCATGCTGTTCAAGAGAGTTTCAGATACACCCGGCTTTGGATCAGC 2287  
2312 QY TCTCAGGGCTTCATGGGAATTTCAATTTCACTCTCAAGATGAAATTTGAATATGAAAACCAA 2371  
2288 Db TCTCAGGGCTTCATGGGAATTTCAATTTCACTCTCAAGATGAAATTTGAATATGAAAACCAA 2347  
2372 QY AAAGGCTG-----GAAGAGAGGAGGACCTTGGAATGTGCTTACATTTGAAAGATCTTCTTT 2425  
2348 Db AGAGGCTGCGCAGAAAGAGGAGGAGAGATTTTGAACGCTGCTGACGTTTGAAGACCTCTTT 2407  
2426 QY GCTTTGCAATCAAGTTGCCAAAGGAATGGAATTTCTGGAATTTAAGTCTGTGTTTACA 2485  
2408 Db GCTTTGCGTACCAAGTTGCCAAAGGAATGGAATTTCTGGAATTTCAAGTCTGTGTTTACA 2467  
2486 QY GAGACCTGGCCCGCAGGAAGTCTCTCACCCAGGGAAGGTGGAAGATGTGACT 2545  
2468 Db GAGACCTGGCCCGCAGGAAGTGTGTGTGTCACCCAGGGAAGGTGGAAGATGTGACT 2527  
2546 QY TTGGATTTGGCTCGAGATATCATGATGATTTCCAACTATGTTGTGAGGGGCAATGCCCGTC 2605  
2528 Db TTGGATTTGGCCCGCAGACATCTGAGCGACTCCAGCTACGTGCTGAGGGCAACGCGCGGC 2587  
2606 QY TGCTGTAAATAGATGGCCCGCAGAAAGCTGTTTGAAGGCAATCTACACCAATTAAGAGTG 2665  
2588 Db TGCCGGTGAATGGATGGACCGCCAGAGCTTATTTTGAAGGCAATCTACCAATCAAGAGTG 2647  
2666 QY ATCTCTGTCATATGGAATATTTCTGTTGGGAATCTTCTCAGTGTGTTGATCTCTTACC 2725  
2648 Db ACCTCTGGTCTTACGGCATCTCTCTCTGGGAGATATTTTCTCAGTGTGTTGATCTCTTACC 2707  
2726 QY CTGGCATTCGGTGTGATGCTAACTTCTACAAAGTGAATTCAAAGTGAATTTAAATGATC 2785

2708 Db CTGGCATTCCTGTGCGCGTAACTTCTTATAAAGTGAATTCAGAGTGAATTTAAATGGAGC 2767  
2786 QY AGCCAATTTATGCTACAGAGAAATATACATTAATGCAATCCTGCTGGGCTTTTGACT 2845  
2768 Db AGCCAATTTATGCTACAGAGAAATATACATTTGTAATGCAATCCTGCTGGGCTTTTGACT 2827  
2846 QY CAAAGAAACGGGCATCTTCCCTAAATTTGAATTTGTTTATGAGTGTGAGTGGGAGATG 2905  
2828 Db CAAAGAAACGGGCATCTTCCCTAAACCTGACTTCAATTTTATGAGTGTGAGTGGGAGAG 2887  
2906 QY CAAAGAAACGGATGTATCAGATGTGGATGGCGCTGTTTCGGAATGCTCTCACACCTACC 2965  
2888 Db CAAAGAAAG-----CATGTATCAGAAATTCATTCATCTCTACC 2924  
2966 QY AAAACAGGCGCATCTTTCAGCAGAGAGATGGAATTTGGGGTACTCTCTCCGCGAGGCTCAGG 3025  
2925 Db AAAACAGGCGCGCTCTCAGCAGAGAGGCGGCTCAGAGC-CCAGTCGCCACAGCGCCAGG 2983  
3026 QY TCGAAGATTTCTGAGAGGAAATTTAGTTTAAAGGATCTTCACTCCCTCCACCTATCCCTAA 3085  
2984 Db T-GAAGATTTACAGAGAAAGATTTAGCGAGGAGGCTTTGGACCCCGC-----CACCTAG 3038  
3086 QY CAGGCTGTAGATTACCAAAACAAAGATTAATTTTATCATCACTAAAGAAATCTATTATCAAC 3145  
3039 Db CAGGCTGTAGACCGCAGAGCCCAAGATTTAGCTCGCCTCT--GAGGAAGCGCCCTACAGCG 3096  
3146 QY TCCTGCTTCCACAGACTTTTCTCTAGAAGCGCTCTGCGTTTACTCTTTTCAAAAGGA 3205  
3097 Db CTTGCTGCTGCTGAGATTTTCTCTAGATGCTGTCTGCCATTACTC-----CAAAGTGA 3149  
3206 QY CTTTGTAAAATCAAATCATCTGTCAAGGCGAGGAGGCTGATAATGAAGCTTTATTG 3265  
3150 Db CTTTGTAAAATCAAACCTCTCTCGCACAGCGGAGAGCCCAATTAATGAGACTTTGTTGG 3209  
3266 QY GAGCATTTGATCTGCATCCAAAGCGCTTCTCAGCGCGCTTGAGTGAATTTGTACCTGAAG 3325  
3210 Db TGAGCGCGCTTACCTTGGGGGCTTTTCCAGAGCTTGAGGGAAAGCCATGTATCTGAAA 3269  
3326 QY TACAGTATATCTTGTAAATACATAAAACAAA-----AGCATTTTGTCTAAGGAGAAAGC 3378  
3270 Db TATAGTATATCTTGTAAATACGTGAACAAACCAACCCGTTTGTCTAAGGAAAGC 3329  
3379 QY TAATATGATTTT---AAGTCTATGTTTAAATAATATGTAATTTTTCAGCTATTAG 3436  
3330 Db TAAATATGATTTTAAATAATCTATGTTTAAATACTATGTAACCTTTTCTATCTATTAG 3389  
3437 QY TGATATATTTTATGGTGGGAATTAATTTCTACTACAGAAAATAAAAAAAAAAAAAA 3496  
3390 Db TGATATATTTTATGGATGGAATAAACTTTCTACTGTAATAAAAAAAAAAAAAA 3449  
3497 QY AAAA 3500  
3450 Db AAAA 3453

Search completed: August 28, 2003, 01:02:16  
Job time : 16331.9 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 15:17:42 ; Search time 8449.93 Seconds  
(without alignments)  
10069.907 Million cell updates/sec

Title: US-09-919-408a-3

Perfect score: 3501

Sequence: 1 CGAGGCGCATCCGAGGCT.....AAAAAAAAAAAAAAAAAAAA 3501

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estopl.*
7: em_estro.*
8: em_htc.*
9: gb_estcl.*
10: gb_estc2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_man.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_fod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*
29: gb_gss2.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3299.4	94.2	3647	11	BC036028
2	846.4	24.2	3761	11	AK045865
3	650.6	18.6	714	12	BI461248
4	644.8	18.4	648	13	BU681356
C					BC036028 Homo sapi AK045865 Mus muscu BI461248 603206574 BU681356 UI-CF-EC1

5	509.6	14.6	621	12	BI360262
6	486.2	13.9	765	9	AI323253
7	440	12.6	604	10	BF523018
8	437.8	12.5	460	9	BF523018
9	402.6	11.5	618	9	AA120050
10	400.4	11.4	455	12	BM147250
11	389.6	11.1	507	12	BM484050
12	382	10.9	510	14	CB220941
13	367.4	10.5	492	13	EX279530
14	358.4	10.2	500	13	EU696796
15	341.2	9.7	3616	11	AK080326
16	340.4	9.7	3312	11	AK004947
17	340.4	9.7	3687	11	AK079247
18	337.2	9.6	3822	11	EC050024
19	334.2	9.5	3687	11	AK076215
20	325.6	9.3	445	10	BE651447
21	318.6	9.1	3883	11	AK046795
22	318.6	9.1	5174	11	BC026713
23	314.4	9.0	522	10	BE644407
24	307.6	8.8	3887	11	AK047010
25	306.8	8.8	597	13	BU695104
26	305.4	8.7	326	12	BM144393
27	299.6	8.6	646	9	AW772610
28	291.6	8.3	658	14	BY727686
29	283.4	8.1	659	9	AW057705
30	274.8	7.8	568	10	BE679272
31	249.2	7.1	515	13	BQ556177
32	242.6	6.9	451	14	CB787035
33	241	6.9	332	14	CB605696
34	240.4	6.9	394	10	EG145808
35	238.2	6.8	724	14	BY764070
36	237.2	6.8	640	14	BY749257
37	236.8	6.8	449	13	BY545844
38	236.6	6.8	660	10	BE047675
39	235.6	6.7	939	13	EX371699
40	234.8	6.7	964	14	CD301880
41	230.8	6.6	779	14	CA511669
42	230.4	6.6	276	10	BF176247
43	228.4	6.5	739	10	BF159205
44	228	6.5	480	9	AI150354
45	227.2	6.5	651	14	CA334084

#### ALIGNMENTS

RESULT 1  
BC036028  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

BC036028 Homo sapiens, clone IMAGE:5272266, mRNA.  
BC036028 Homo sapiens (human)  
BC036028.1 GI:23273564  
HTC.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Strausberg,R.  
Direct Submission  
Submitted (11-JUL-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki  
Toshiyuki and Piero Carninci (RIKEN)  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
Center, Stanford University School of Medicine, Stanford, CA 94305

REMARK COMMENT

Web site: <http://www.shgc.stanford.edu>  
Contact: (Dickson, Mark) mcd@paxil.stanford.edu  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAC Plate: 48 Row: d Column: 24  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4758395  
This clone has the following problem: frame shifted.

## FEATURES

source  
1..3647  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:527266"  
/tissue\_type="Testis"  
/clone\_lib="NIH\_MGC\_97"  
/lab\_host="DH10B"  
/note="Vector: pBluescript"  
BASE COUNT 1114 a 735 c 829 g 969 t  
ORIGIN

Query Match 94.2%; Score 3299.4; DB 11; Length 3647;  
Best Local Similarity 95.9%; Pred. No. 0;  
Matches 3473; Conservative 0; Mismatches 16; Indels 134; Gaps 3;

Qy	1	CGAGGGCGCATCCAGGGCTGGCGGGCGG-CCTGGGGGACCCGGGCTCGGAGGCCAT	59
Db	25	CCAGGGCGCATCGAGGGCTGGCGGGCGGCTGGGGACCCGGGCTCGGAGGCCAT	84
Qy	60	GCCGGCGTTGGCGGCGAGCGGGCGACCGTCCCGCTGCTGCTGTTTCTGCAATGAT	119
Db	85	GCCGGCGTTGGCGGCGAGCGGGCGAGCTGCCCTGCTGCTGTTTCTGCAATGAT	144
Qy	120	ATTGGGACTATTAACAATCAAGATCTGCCCTGATCAAGTGTCTTTTAATCAATATAA	179
Db	145	ATTGGGACTATTAACAATCAAGATCTGCCCTGATCAAGTGTCTTTTAATCAATATAA	204
Qy	180	GAACATGATTCAGTGGGAGTCAATCATATCCATGATCAGATCCCGGA	239
Db	205	GAACATGATTCAGTGGGAGTCAATCATATCCATGATCAGATCCCGGA	264
Qy	240	AGACCTCGGTGTCGTTGACGCCAGAGCTCAGGACAGTGTACGAAGTCCGGTGT	299
Db	265	AGACCTCGGTGTCGTTGACGCCAGAGCTCAGGACAGTGTACGAAGTCCGGTGT	324
Qy	300	GGAGTGGATGTATCTGCTCCATCACA CTGCAAGTGTGTCGATGCCCGGAGACAT	359
Db	325	GGAGTGGATGTATCTGCTCCATCACA CTGCAAGTGTGTCGATGCCCGGAGACAT	384
Qy	360	TTCTGTCTCGGTCTTTAAGCAGAGTCCCTGAATGCCAGCACATTTGATTACA	419
Db	385	TTCTGTCTCGGTCTTTAAGCAGAGTCCCTGAATGCCAGCACATTTGATTACA	444
Qy	420	AAACAGAGGAGTGTCTTCCATGTTGCAATTTGCAAAATGACAGAAACCCAGCTGAGAATA	479
Db	445	AAACAGAGGAGTGTCTTCCATGTTGCAATTTGCAAAATGACAGAAACCCAGCTGAGAATA	504
Qy	480	CCTACTTTTATTACAGAGTGAAGCTACCAATTACCAATATTGTTACAGTGAATTAAG	539
Db	505	CCTACTTTTATTACAGAGTGAAGCTACCAATTACCAATATTGTTACAGTGAATTAAG	564
Qy	540	AAATACCTGTTTACATTAAGAAGACCTTACTTTAGAAAAATGAAAAACCGAGCGC	599
Db	565	AAATACCTGTTTACATTAAGAAGACCTTACTTTAGAAAAATGAAAAACCGAGCGC	624
Qy	600	CCTGGTCTGATATCTCAGAGCGTTCCAGAGCCGATCGTGAATGGGTGCTTCGGATTC	659
Db	625	CCTGGTCTGATATCTCAGAGCGTTCCAGAGCCGATCGTGAATGGGTGCTTCGGATTC	684
Qy	660	ACAGGGGAAAGCTGTAAGAAAGAGTCCAGCTGTTGTTAAAAAGGAGGAAAAAGTGCT	719

Db	685	ACAGGGGAAAGCTGTAAGAAAGAGTCCAGCTGTTGTTAAAAAGGAGGAAAAAGTGCT	744
Qy	720	TCATGAATTTATTCGGACGGACATAAGGTGCTGTGCCAGAAATGAATCGGCGAGGAATG	779
Db	745	TCATGAATTTATTCGGACGGACATAAGGTGCTGTGCCAGAAATGAATCGGCGAGGAATG	804
Qy	780	CACGAGCTGTTCAATATAGATCTAAATCAAACTCCTCAGACCAATTTGCCACAATATT	839
Db	805	CACGAGCTGTTCAATATAGATCTAAATCAAACTCCTCAGACCAATTTGCCACAATATT	864
Qy	840	TCITTAAGTAGGGAAACCTTATGAGTAAGGTGCAAGCTGTTTATGTAACCATGAT	899
Db	865	TCITTAAGTAGGGAAACCTTATGAGTAAGGTGCAAGCTGTTTATGTAACCATGAT	924
Qy	900	CGGCTCACCTGGGAAATTAGAAAAACAAAGCACTCGAGGAGGCAACTACTTTGAGATGAG	959
Db	925	CGGCTCACCTGGGAAATTAGAAAAACAAAGCACTCGAGGAGGCAACTACTTTGAGATGAG	984
Qy	960	TACCTATTCAACAAACAGAACTATGATACGGAATCTGTTTCTTTTGTATCATCATGCGC	1019
Db	985	TACCTATTCAACAAACAGAACTATGATACGGAATCTGTTTCTTTTGTATCATCATGCGC	1044
Qy	1020	AGAAACGACACCGGATACACTGTTTCTCTTCAAAGCATCCAGTCAATCAGCTTT	1079
Db	1045	AGAAACGACACCGGATACACTGTTTCTCTTCAAAGCATCCAGTCAATCAGCTTT	1104
Qy	1080	GGTTTACCATCGTAGAAGGATTTATAAATGCTACCAATTTCAAGTGAAGATTTATGAAT	1139
Db	1105	GGTTTACCATCGTAGAAGGATTTATAAATGCTACCAATTTCAAGTGAAGATTTATGAAT	1164
Qy	1140	TGACCAATATGAAGATTTTGTCTGTGCTGAGGTTTAAAGCCTACCCACAAATCAGATG	1199
Db	1165	TGACCAATATGAAGATTTTGTCTGTGCTGAGGTTTAAAGCCTACCCACAAATCAGATG	1224
Qy	1200	TACGTGACCTTCTCTCGAAATCATTTCTTGTGAGCAAAAGGCTTTGATAACGATA	1259
Db	1225	TACGTGACCTTCTCTCGAAATCATTTCTTGTGAGCAAAAGGCTTTGATAACGATA	1284
Qy	1260	CAGCATATCAAGTTTGCATATAGCAACCCAGGAGAAATATATTTCCATGCAGA	1319
Db	1285	CAGCATATCAAGTTTGCATATAGCAACCCAGGAGAAATATATTTCCATGCAGA	1344
Qy	1320	AAATGATGATGCCAAATTTTACCAAAATGTTTCACTGAAATATAGAAAGAAACCTCAAGT	1379
Db	1345	AAATGATGATGCCAAATTTTACCAAAATGTTTCACTGAAATATAGAAAGAAACCTCAAGT	1404
Qy	1380	GCTCGCAAGAGCATCGGCAAGTCAAGGCTGCTGTTTCTCGGATGGATACCCATACCATC	1439
Db	1405	GCTCGCAAGAGCATCGGCAAGTCAAGGCTGCTGTTTCTCGGATGGATACCCATACCATC	1464
Qy	1440	TTGGACCTGGAAGAGTGTTCAGCAAGTCTCCCACTGCAAGAGGATCAGAGG	1499
Db	1465	TTGGACCTGGAAGAGTGTTCAGCAAGTCTCCCACTGCAAGAGGATCAGAGG	1524
Qy	1500	AGTCTGGAATAGAAAGGCTAACAGAAAAGTGTGGAAGAGTGTGCGAGCTACTCT	1559
Db	1525	AGTCTGGAATAGAAAGGCTAACAGAAAAGTGTGGAAGAGTGTGCGAGCTACTCT	1584
Qy	1560	AAACATGAGTGAAGCCATAAAAGGTTCTGGTCAAGTGTGTCATACAAATCCCTTGG	1619
Db	1585	AAACATGAGTGAAGCCATAAAAGGTTCTGGTCAAGTGTGTCATACAAATCCCTTGG	1644
Qy	1620	CACATCTTGTGAGAGCATCTTTTAAACTCTCAGGCGCTTCCCTTTTCATCCAGACAA	1679
Db	1645	CACATCTTGTGAGAGCATCTTTTAAACTCTCAGGCGCTTCCCTTTTCATCCAGACAA	1704
Qy	1680	CATCTCATTTATGCAACAAATGCTGTTGCTCCTCTTCAATGCTGTTTTAAACCTGCT	1739
Db	1705	CATCTCATTTATGCAACAAATGCTGTTGCTCCTCTTCAATGCTGTTTTAAACCTGCT	1764
Qy	1740	AAATTTGTCAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGT	1799

Db 1765 AATTGTGACAAAGTACAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGT 1824  
 Qy 1800 GACCGGCTCCTCAGATAATGAGTACTTCTAGTGTGATTTTACAGAAAT-ATGAAATATGTC 1858  
 Db 1825 GACCGGCTCCTCAGATAATGAGTACTTCTAGTGTGATTTTACAGAAATATGATGTC 1884  
 Qy 1859 TCAAAATGGGAGTTTCCCAAGAGAAAATTTAGAGTTTGGGAGGTACTAGGATCAGGTGCTT 1918  
 Db 1885 TCAAAATGGGAGTTTCCCAAGAGAAAATTTAGAGTTTGGGAGGTACTAGGATCAGGTGCTT 1944  
 Qy 1919 TTGAAAAGTGAAGAACGCAACAGCTTATGAGAAATTTAGCAAAACAGGAGTCTCAATCAGG 1978  
 Db 1945 TTGAAAAGTGAAGAACGCAACAGCTTATGAGAAATTTAGCAAAACAGGAGTCTCAATCAGG 2004  
 Qy 1979 TTGCGGTCAAAATGCTCAAAAGAAAAGCAGACAGCTCTGAAAGAGAGCACTCATGTGAC 2038  
 Db 2005 TTGCGGTCAAAATGCTCAAAAGAAAAGCAGACAGCTCTGAAAGAGAGCACTCATGTGAC 2064  
 Qy 2039 AACTCAAGATGATGACCCAGCTGGGAAAGCCACGAGAAATATTGTGAACCTGCTGGGGCGT 2098  
 Db 2065 AACTCAAGATGATGACCCAGCTGGGAAAGCCACGAGAAATATTGTGAACCTGCTGGGGCGT 2124  
 Qy 2099 GCACACTGTGAGGACCAATTTACTTTGATTTTGAATATCTGTGCTATGATCTTCTCA 2158  
 Db 2125 GCACACTGTGAGGACCAATTTACTTTGATTTTGAATATCTGTGCTATGATCTTCTCA 2184  
 Qy 2159 ACTATCTAAGAGTAAAGAGAAAATTTTACAGGACTTGGACAGAGATTTTCAAGGAAC 2218  
 Db 2185 ACTATCTAAGAGTAAAGAGAAAATTTTACAGGACTTGGACAGAGATTTTCAAGGAAC 2244  
 Qy 2219 ACAATTTCACTTTTACCCACATTTTCCAAATCACATCCAAATTTCCAG- 2264  
 Db 2245 ACAATTTCACTTTTACCCACATTTTCCAAATCACATCCAAATTTCCAGTAAAGAAATGAG 2304  
 Qy 2265 ----- 2264  
 Db 2305 CTTTACAAAGCAACTGGAAGAAAAGGATGTTGAAACGCTTACGGGACTCTCGGAA 2364  
 Qy 2265 ----- 2264  
 Db 2365 GATCTGTATTATGTGAGGAAAAGTGGGCTGAGCTCAGAAACCAAGAAATGAGATCGATCA 2424  
 Qy 2267 TGCGTGGTTCAAGAGAGTTTCAAGTACACCGGACTCGGATCAATCTCAGGGCTTCATG 2326  
 Db 2425 TGCGTGGTTCAAGAGAGTTTCAAGTACACCGGACTCGGATCAATCTCAGGGCTTCATG 2484  
 Qy 2327 GGAATTCATTTCACTCTGAAAGATGAAATTTGAATATGAAACCAAAAGGCTGGAAGAAG 2386  
 Db 2485 GGAATTCATTTCACTCTGAAAGATGAAATTTGAATATGAAACCAAAAGGCTGGAAGAAG 2544  
 Qy 2387 AGGAGACTTGAATGCTTACATTTGAAAGATCTTCTTTGCTTTGCAATATCAAGTTGCCA 2446  
 Db 2545 AGGAGACTTGAATGCTTACATTTGAAAGATCTTCTTTGCTTTGCAATATCAAGTTGCCA 2604  
 Qy 2447 AAGGAATGGAATTTCTGGAATTTAAGTCTGTTGTTTCAAGAGACTCTGGCCCGCAGGAACG 2506  
 Db 2605 AAGGAATGGAATTTCTGGAATTTAAGTCTGTTGTTTCAAGAGACTCTGGCCCGCAGGAACG 2664  
 Qy 2507 TGCTTTGACCCACCGGAAAAGTGGTGAAGATATGTAATTTGGAATTTGCTCGAGATATCA 2566  
 Db 2665 TGCTTTGACCCACCGGAAAAGTGGTGAAGATATGTAATTTGGAATTTGCTCGAGATATCA 2724  
 Qy 2567 TGAGTGAATCCAACTATGTTTGTGAGGGGCAATGCCCGTCTGCTGTGAAAATGGAATGCC 2626  
 Db 2725 TGAGTGAATCCAACTATGTTTGTGAGGGGCAATGCCCGTCTGCTGTGAAAATGGAATGCC 2784  
 Qy 2627 CCGAAGCCCTGTTTGAAGGATCTACACATTAAGAGTATGTTGCTGATATGGAATAT 2686  
 Db 2785 CCGAAGCCCTGTTTGAAGGATCTACACATTAAGAGTATGTTGCTGATATGGAATAT 2844  
 Qy 2687 TACTGTGGGAAATCTTCTCACTGTTGTAATCTTACCTGCAATTCAGGATTCGCTGATGCTA 2746  
 Db 2845 TACTGTGGGAAATCTTCTCACTGTTGTAATCTTACCTGCAATTCAGGATTCGCTGATGCTA 2904

Qy 2747 ACTTCTACAAAGTATTTCAAAATGGAATTTAAATGGAATGGAATTTTATGCTACAGAG 2806  
 Db 2905 ACTTCTACAAAGTATTTCAAAATGGAATTTAAATGGAATGGAATTTTATGCTACAGAG 2964  
 Qy 2807 AAATATATCATTTAATGCAATCTCTGCTGGGCTTTGACTCAAGGAAACGGCCATCTTCC 2866  
 Db 2965 AAATATATCATTTAATGCAATCTCTGCTGGGCTTTGACTCAAGGAAACGGCCATCTTCC 3024  
 Qy 2867 CTAAATTTGACTTCTGTTTATAGGATGTCAGCTGGCAGATGCAAGAAAGCGATGATCA 2926  
 Db 3025 CTAAATTTGACTTCTGTTTATAGGATGTCAGCTGGCAGATGCAAGAAAGCGATGATCA 3084  
 Qy 2927 ATGTGAGTGGCGGCTTTTCGGAATGCTCCTCACACTACCAAAACAGGCGACTTTTCC 2986  
 Db 3085 ATGTGAGTGGCGGCTTTTCGGAATGCTCCTCACACTACCAAAACAGGCGACTTTTCC 3144  
 Qy 2987 GAGAGATGGAATTTTGGGGCTACTCTCTCCGAGGCTCAGGTGCAAGATTTCTGAGGAAACA 3046  
 Db 3145 GAGAGATGGAATTTTGGGGCTACTCTCTCCGAGGCTCAGGTGCAAGATTTCTGAGGAAACA 3204  
 Qy 3047 ATTTAGTTTAAAGACTTTCATCCCTCACCTATCCCTAAACAGGCTGTAGATTTACCAAAAC 3106  
 Db 3205 ATTTAGTTTAAAGACTTTCATCCCTCACCTATCCCTAAACAGGCTGTAGATTTACCAAAAC 3264  
 Qy 3107 AAGATTAATTTTCACTAAAGAAATCTATTAATCAACTGCTGCTTCCAGACTTTTTC 3166  
 Db 3265 AAGATTAATTTTCACTAAAGAAATCTATTAATCAACTGCTGCTTCCAGACTTTTTC 3324  
 Qy 3167 TCTAGAAAGCGCTGCTGCGTTTACTCTCTTTTCAAGGAGCTTTTGTAAAAATCAATATC 3226  
 Db 3325 TCTAGAAAGCGCTGCTGCGTTTACTCTCTTTTCAAGGAGCTTTTGTAAAAATCAATATC 3384  
 Qy 3227 CTGTCAAGGCGGAGGAGCTGATTAATGAATTTATTGAGGACTTGTATCTGCAATCCAAG 3286  
 Db 3385 CTGTCAAGGCGGAGGAGCTGATTAATGAATTTATTGAGGACTTGTATCTGCAATCCAAG 3444  
 Qy 3287 GCCTTCTCAGGCGGCTGAGTGAATTTGTAAGTCTGTAAGTACAGTATATTTCTGTAATA 3346  
 Db 3445 GCCTTCTCAGGCGGCTGAGTGAATTTGTAAGTCTGTAAGTACAGTATATTTCTGTAATA 3504  
 Qy 3347 CATAAACCAAGCAATTTTCTAAGGAGAGCTTAATATGATTTTAAAGTCTATGTTTAA 3406  
 Db 3505 CATAAACCAAGCAATTTTCTAAGGAGAGCTTAATATGATTTTAAAGTCTATGTTTAA 3564  
 Qy 3407 AAATAATATCTAAATTTTTCAGCTTATTAGTATATATTTTATGGTGGGAAATAAAAAT 3466  
 Db 3565 AAATAATATCTAAATTTTTCAGCTTATTAGTATATATTTTATGGTGGGAAATAAAAAT 3624  
 Qy 3467 CTACTACAGAAAAAATAAAAAA 3489  
 Db 3625 CTACTACAGAAAAAATAAAAAA 3647

RESULT 2  
 AK045865  
 LOCUS  
 DEFINITION  
 AK045865 3761 bp mRNA linear HTC 05-DEC-2002  
 Mus musculus adult male corpora quadrigemina cDNA, RIKEN  
 full-length enriched library, clone:B20315G04 product:FMS-like  
 tyrosine kinase 3, full insert sequence.  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Mus musculus  
 Mus musculus (house mouse)  
 HTG; CAP trapper.  
 AK045865.1 GI:26337656  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCES  
 1  
 AUTHORS  
 TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning  
 Meth. Enzymol. 303, 19-44 (1999)  
 99279253  
 10349636

REFERENCE  
AUTHORS

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL  
MEDLINE  
PUBMED

20499374  
11042159

REFERENCE  
AUTHORS

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, F., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipillar sequencer  
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL  
MEDLINE  
PUBMED

20530913  
11076861

REFERENCE  
AUTHORS

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Harada, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L.M., Stauber, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Balzarrelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., De Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seva, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)

JOURNAL  
MEDLINE  
PUBMED

21085660  
11217851

REFERENCE  
AUTHORS

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
MEDLINE  
PUBMED

11217851  
(bases 1 to 3761)

REFERENCE  
AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozawa, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.  
Direct Submission

## JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]

## COMMENT

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/

## FEATURES

## source

Location/Qualifiers  
1. .3761  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_DB:B230315G04"  
/db\_xref="taxon:10090"  
/c1one="B230315G04"  
/sex="male"  
/tissue\_type="corpora quadrigemina"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
48. .1268  
/note="unnamed protein product; FMS-like tyrosine kinase 3 (MGD|MG1:95559, GB|NM\_010229, evidence: BLASTN, 99%, match=1238)  
putative"  
/codon\_start=1  
/protein\_id="BAC32514.1"  
/db\_xref="GI:26337657"  
/translation="MRALORSRRRLVLLVLSVILETVTNDLPVTKCVLI SHENN GSSAKRPSRVMVRGSPEDLCAPRROSEGTVEAATVEUABSSITLQVLAIPGDL SCWLVFHSLSGQPHFDLQNRGIVSNAILNVTQAGEYLLHQSAAVTLVFTVN VRTQVLYLRRPYFRKNQDLALCISEGPEPTVWLCSHRESKEEPAVVRKE EKVYHELFGTDIRCCARNALDRCTKLFITDLNQAQOSTLPQLFKVGEPLWIRKAI HVNHGFLTWELDEKALEGSYPFMYSTNRTMIRILLAFVSVGRNDYGYTCSSS KHPQSALVTILKGFNATSSQEEYIDPVEKFCFVRFKAYPRIRCTWTFQSASF CEORGLEDGRLSH"

## CDS

BASE COUNT 902 a 950 c 948 g 961 t  
ORIGIN  
Query Match 24.2%; Score 846.4; DB 11; Length 3761;  
Best Local Similarity 80.6%; Pred. No. 2e-97;  
Matches 1015; Conservative 0; Mismatches 241; Indels 4; Gaps 2;  
17 GGCTGGGCGCGCGCTGGGGGACCCCGGCTCCGAGGCCATCGCGCGTTGGC 73  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 66  
8 GGCTGGGCGCGCGCGCT-GGCTACCGCGCTCCGAGGCCATCGCGCGTTGGC 66  
Qy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 66  
74 GCAGCGGGGACCGTCCGCTGCTGCTGTTGTTTCTGCAATCATATTTGGGACTATA 133  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 126  
67 GCAGCGCGCGCGCTGCTGCTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 126  
Qy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 126  
134 CAAATCAAGATCTGCCGTGATCAAGTGTTGTTTAAATCAATCAATCAATCAATCAAT 193  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 186  
127 CAAACCAAGACCTGCCGTGATCAAGTGTTGTTTAAATCAAGTCAAGTCAAGTCAAGT 186  
Qy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 186  
194 CAGTGGGGAAGTCAATCATCATATATCCCATGGTATCAGAAATCCCGGAAGACCTCGG 253  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 246  
187 CAGCGGGAAGCCATCATCTGATCCGATGGTGGGAGGATCCCAAGAGACCTCCAGTGTG 246  
Qy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 246  
254 CGTTGAGACCCGAGCTCAGGAGCAGTGTACGAAGTGGCGTGTGGAAAGTGAATGAT 313  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 306  
247 CCCGAGGCGCGAGTGAAGGAGCGGTATATGAAGCGGCCCGTGGAGTGGCCGAGT 306  
Qy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 373  
314 CTGCTTCATCAGTGCAGTGCATGATGATGATGATGATGATGATGATGATGATGATGAT 373  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 366  
307 CTGGGTCCATCACCTGCAAGTGCAGTCCGCCACCCAGGAGGACCTTTCCTGCTCTGG 366  
Qy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 433  
374 TCTTTAGACAGCTCCCTGAAATGCCAGCCACATTTTGAATTTTCAAAACAGAGAGTGT 433  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 426  
367 TCTTTAGACAGCTCCCTGCGGTGCCAGCGGCACTTTGATTTTACAAAACAGAGAGTGT 426  
Qy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 493  
434 TTTCCATGGTCAATTTTGAATGACAGAAACCCAGCTGGAGATACCTACTTTTATTC 493  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 486  
427 TTTCCATGGTCAATTTTGAATGACAGAAACCCAGCTGGAGATACCTACTTTTATTC 486



QY	494	AGAGTGAAGCTAACAAATTACAAATATTGTTTACAGTGAGTATAAGAAATAACCTCGCTTT	553
DB	487	AGAGCGAAGCGCCAACTACACAGTACTGTTTACAGTGAATGAAGAGATACACAGCTGT	546
QY	554	ACACATTAAGAGAGACCTTACTTTAGAAAAATGGAAAAACGAGACGCCCTGCTGCATAT	613
DB	547	ACGTGCTAAGAGAGACCTTACTTTTAGGAAGATGGAAAAACGAGACGCACTGCTCTGCATCT	606
QY	614	CTGAGAGCGTTCCAGAGCCGATCGTGGAAATGGGTCTTTGCGATTACAGGGGGAAAGCT	673
DB	607	CCGAGGGTGTCCAGAGCCCACTGTGGATGGGTGCTCTGCAGCTCCACAGGGAAAGCT	666
QY	674	GTAAGAGAAAGTCCAGCTGTTGTTTAAAAAGGAGGAAAAAGTGCTTCATGAATATTG	733
DB	667	GTAAGAGAGAAAGGCCCTGCTGTTGTGAGAAAGGAGAAAGGTACTTTCATGAGTTGTTG	726
QY	734	GGACGGAATAAAGTGTGTCGCCAGAAATGAACCTGGGCAGAGGAATGCACAGGCTGTTC	793
DB	727	GAAACAGACATCAGATGCTGTGCTTGAATGCACTGGGCCGGAATGCACCAAGCTGTTC	786
QY	794	CAATPAGATCTAAATCAAACTCCTCAGAGCACAATTGCCACAATTATTTCTTAAAGTAGGGG	853
DB	787	CCATPAGATCAAAACAGGCTCCTCAGAGCACACTGCCCACTATTCTGAAAGTGGGG	846
QY	854	AACCCCTATTGATTAAGGTGCAAAAGCTGTTTCAATGTGAACCAATGGAATCGGGCTACCTGGG	913
DB	847	AACCCCTTTGTGATCAGGTGTAAAGCCCATCCACTGTGAACCATGGATTGGGCTCACCTGGG	906
QY	914	AATTAGAAAAACAAGACTCGAGAGGGCACTACTTTGAGATGAGTACCTATTTCACAA	973
DB	907	AGCTGGAAGACAAGACCCCTGGAGAGGCACTACTTTGATGATGAGTACCTACTTCCA	966
QY	974	ACGAACTATGATACGGATTCTGTTTGTCTTTTGTATCATCAGTGGCAAGAAAACGACACCG	1033
DB	967	ACAGGACCATGATTCGGATTCTCTTGGCTTTTGTCTTCCGTGGGAAGGAAACGACACCG	1026
QY	1034	GATACTACATTTGTTTCTCTTTCAAAGCATCCAGTCAATCAGCTTTTGGTTACCATCGTAG	1093
DB	1027	GATATTACAGCTGTCTTCTCTCAAAGCAACCCAGCCAGTCAGCGTTTGGTGAACCATCTAG	1086
QY	1094	GAAAGGATTTATAAATGCTACCAATTCAGTGAAGATTATGAATTTGACCAATATGAAG	1153
DB	1087	AAAAGGGTTTATAACGCTACCACTCGCAAGAAAGAGTATGAATTTGACCCCGTACGAA	1146
QY	1154	AGTTTTTGTCTTGTGTCAGGTTTAAAGCCTACCCACAAATCAGATGTACCGTGACCTTCT	1213
DB	1147	AGTTCTGCTTCTCAGTCAGGTTTAAAGGTTACCCAGAAATCCGATGCACTGATCTTCT	1206
QY	1214	CTCGAAATCATTTCTTGTGAGCAAAAGGGTCTTTGATACCGGATACAGCATATCCAAGT	1273
DB	1207	CTCAAGCCTCATTTCTTTGTAACAGAGAGGCTTGGAGGATGGGTACAGACTCTCTCAT	1266

RESULT 3  
 BI461248  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

714 bp mRNA linear EST 21-AUG-2001  
 603206574F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:5272366 5',  
 mRNA sequence.  
 BI461248  
 BI461248.1 GI:15251904  
 EST.  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 714)  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.



Qy 600 CCTGGTCTGCATATCTAGAGCGTTCCAGAGCCGATCGTGAATGGGTGCTTTGGGATTC 659  
 Db 630 CCTGGTCTGCATATCTAGAGCGTTCCAGAGCCGATCGTGAATGGGTGCTTTGGGATTC 689  
 Qy 660 ACAGGGGAAAGCTGTAAGAGAA 684  
 Db 690 ACAGGGGAAAGCTGTAAGAGAA 714

RESULT 4  
 BU681356/c 648 bp mRNA linear EST 07-OCT-2002  
 LOCUS UI-CF-EC1-abv-d-14-0-UI.s1 UI-CF-EC1 Homo sapiens cDNA clone  
 DEFINITION UI-CF-EC1-abv-d-14-0-UI 3', mRNA sequence.

ACCESSION BU681356  
 VERSION BU681356.1 GI:23531192  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 648)  
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene  
 discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477

PUBMED 8889548  
 COMMENT Contact: McCray, PB  
 McCray Lab

University of Iowa  
 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA

Tel: 319 356 4866  
 Fax: 319 356 7171

Email: paul-mccray@uiowa.edu  
 Tissue Procurement: Dr. M. J. Welsh, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com) or from Open Biosystems  
 (www.openbiosystems.com).

Seq primer: M13 FORWARD  
 POLYA=Yes.

FEATURES  
 source Location/Qualifiers  
 1..648 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-CF-EC1-abv-d-14-0-UI"  
 /tissue\_type="Lung"  
 /dev\_stage="Adult and Fetal"  
 /lab\_hosts="DH10B (Life Technologies) (T1 phage resistant)"  
 /clone\_libs="UI-CF-EC1"

/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a  
 modified polylinker; Site 1: EcoR I; Site 2: Not I;  
 UI-CF-EC1 is a normalized cDNA library containing the

following tissue(s): Normal lung from adult and from fetal  
 day 64, day 87, week 19 and week 42. The library was

constructed according to Bonaldo, Lennon and Soares,  
 Genome Research, 6:791-806, 1996. First strand cDNA

synthesis was primed with an oligo-dT primer containing a  
 Not I site. Double stranded cDNA was ligated to an EcoR I

adaptor, digested with Not I, and cloned directionally  
 into pT7T3-Pac vector. The oligonucleotide used to prime

the synthesis of first-strand cDNA contains a library tag  
 sequence that is located between the Not I site and the

(dT)18 tail. The sequence tag for this library is  
 AAGTGCCTAC.

TAG\_LIB=UI-CF-EC1  
 TAG\_TISSUE=Normal Lung Epithelial Cells Tissue nos 369-371  
 and 380-383

TAG\_SEQ=AAGTGCCTAC"

BASE COUNT 196 a 126 c 120 g 206 t  
 ORIGIN

Query Match 18.4%; Score 644.8; DB 13; Length 648;  
 Best Local Similarity 99.7%; Pred. No. 9.6e-72;  
 Matches 646; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2844 CTCAAGGAAACGGCCATCTTCCCTAATTTGACTTCGTTTTAGGATGTGACGTGGCAGA 2903  
 Db 648 CTCAAGGAAACGGCCATCTTCCCTAATTTGACTTCGTTTTAGGATGTGACGTGGCAGA 589

Qy 2904 TGCAGAAAGCCGATGTATCAGAAATGGATGGCCGTGTTTCGGAATGTCTCACACCTA 2963  
 Db 588 TGCAGAAAGCCGATGTATCAGAAATGGATGGCCGTGTTTCGGAATGTCTCACACCTA 529

Qy 2964 CCAAAACAGCGCCTTTTCAGCAGAGAGATGGATTTGGGGCTACTCTCTCCGCAAGGCTCA 3023  
 Db 528 CCAAAACAGCGCCTTTTCAGCAGAGAGATGGATTTGGGGCTACTCTCTCCGCAAGGCTCA 469

Qy 3024 GGTCGAAGATTCGTPAGAGGAACAATTTAGTTTTAAGGACTTCATCCCTCCACCTATCCCT 3083  
 Db 468 GGTCGAAGATTCGTPAGAGGAACAATTTAGTTTTAAGGACTTCATCCCTCCACCTATCCCT 409

Qy 3084 AACGGCTGTAGATTACCAAAAACAGATTAATTTTCATCATCACTAAAAAGAAAATCTATTATCA 3143  
 Db 408 AACGGCTGTAGATTACCAAAAACAGATTAATTTTCATCATCACTAAAAAGAAAATCTATTATCA 349

Qy 3144 ACTGCTGCTTCCACAGACTTTTCTCTAGAAGCCGTCTGCGTTTACTCTTGTGTTTCAAAGG 3203  
 Db 348 ACTGCTGCTTCCACAGACTTTTCTCTAGAAGCCGTCTGCGTTTACTCTTGTGTTTCAAAGG 289

Qy 3204 GACTTTTTGTAATAATCAAAATCATCTGTCAAGAGCAGGAGGAGCTGATAATGAATTTAT 3263  
 Db 288 GACTTTTTGTAATAATCAAAATCATCTGTCAAGAGCAGGAGGAGCTGATAATGAATTTAT 229

Qy 3264 TGGAGCATTTGATCTGCATCCAAAGCCCTTCTCAGCGCCGCTTGTAGTGAATTTGTACCTGA 3323  
 Db 228 TGGAGCATTTGATCTGCATCCAAAGCCCTTCTCAGCGCCGCTTGTAGTGAATTTGTACCTGA 169

Qy 3324 AGTACAGTATATCTTCTGTAATACATAAAACAAAGCATTTTGTCTAGGAGAACTAATA 3383  
 Db 168 AGTACAGTATATCTTCTGTAATACATAAAACAAAGCATTTTGTCTAGGAGAACTAATA 109

Qy 3384 TGATTTTTTAAGTCTATGTTTTTAAATAATATGTAATAATTTTTCAGCTATTAGTGATATA 3443  
 Db 108 TGATTTTTTAAGTCTATGTTTTTAAATAATATGTAATAATTTTTCAGCTATTAGTGATATA 49

Qy 3444 TTTTATCGGTGGGAATAAAAATTTCTACTACAGAAAAAATAAAAAA 3491  
 Db 48 TTTTATCGGTGGGAATAAAAATTTCTACTACAGAAAAAATAAAAAA 1

RESULT 5  
 BI360262 621 bp mRNA linear EST 01-AUG-2001  
 LOCUS BI360262

DEFINITION 387099 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.  
 ACCESSION BI360262  
 VERSION BI360262.1 GI:15056290

KEYWORDS EST.  
 SOURCE Sus scrofa (pig)  
 ORGANISM Sus scrofa

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1 (bases 1 to 621)  
 AUTHORS Fahrenkrug,S.C., Smith,T.P.L.; Freking,B.A., Cho,J., White,J.,  
 Vallet,J., Wise,T., Rohrer,G.A., Pertea,G., Sultana,R., Quackenbush,  
 J. and Keele,J.W.

TITLE Porcine gene discovery by normalized cDNA-library sequencing and  
 EST cluster assembly

JOURNAL Mamm. Genome 13 (8), 475-478 (2002)  
 MEDLINE 22213789  
 PUBMED 12226715

COMMENT Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: smith@mail.marc.usda.gov

Single pass sequencing. Bases called and alt trimmed with phred  
v0.980904.e. Vector identified by cross\_match with the -minscore 18  
and -minmatch 12 options.  
PCR Primers  
FORWARD: AGGAACAGCTATGACCAT  
BACKWARD: GTTTCACGATCAGGAGC  
Plate: 139 row: H column: 2  
Seq primer: ATTAGGTGACACTATAG.

Location/Qualifiers  
1. 621  
/organism="Sus scrofa"  
/mol\_type="mRNA"  
/db\_xref="taxon:9823"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/clone\_lib="MARC 2P1G"  
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;  
Library made from pooled tissue from testis, ovary,  
endometrium, hypothalamus, pituitary, and placenta."

FEATURES

source

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches 551; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

1725 CGTTTAAACCTGCTAATTTGTCACAGTACAAAGCAAAATTTAGGTATGAAGCCAGCT 1784

2 CGTTTAAACCTGCTAATTTGTCACAGTACAAAGCAAAATTTAGGTATGAAGCCAGCT 61

1785 ACAGATGGTACAGGTGACCGGCTCTCAGATAATGAGTACTTCTAGCTTGATTTTCAGAGA 1844

62 GCAGATGGTCCAGGTGACGGGTCCTGGATACGACTACTTCTACATCGACTTCAGAGA 121

1845 ATATGAATATGATCTCAAAATGGGAGTTTCCAGAGAGAAAATTTAGAGTTTGGGAAGTACT 1904

122 ATATGAATATGATCTCAAAATGGGAGTTTCCAGAGAGAAAATTTAGAGTTTGGGAAGTACT 181

1905 AGGATCAGGTGCTTTTGGAAAAGTATGACGCAACAGCTTATGGAATATGACAAAAGG 1964

182 GGGATCGGGCGCTTTTGGAAAAGTATGACGCAACAGCTTATGGAATATGACAAAAGG 241

1965 AGTCTCAATCCAGGTGCGGCTCAAAATGCTGAAAAGAAAAGCAGACAGCTCTGAAAGAGA 2024

242 AGTCTCAATCCAGGTGCGGCTCAAAATGCTGAAAAGAAAAGCAGACAGCTCTGAAAGAGA 301

2025 GGCACATCATCTAGAACTCAGATGATGACCCAGCTGGGAAGCACCAGAGATATTTGTGAA 2084

302 GGCCTCATCTGTAACCTCAAAATGATGACCCAGCTGGGAAGCACCAGAGATATTTGTGAA 361

2085 CTTGCTGGGGGGCTGACACTGTGACAGCAAAATTTACTGATTTTGAATTTGCTGCTA 2144

362 CTTGCTGGGGGGCTGACACTGTGACAGCAAAATTTACTGATTTTGAATTTGCTGCTA 421

2145 TGGTGATCTTCTCAACTATCTAAGAACTAAGAGAAAATTTTCAGAGGCTTGGACAGA 2204

422 CGGTGATCTTCTCAACTATCTAAGAACTAAGAGAAAATTTTCAGAGGCTTGGACCCA 481

2205 GATTTTCAAGGAACACAATTTTCACTTTTCCCACTTTTCCCAATCATCAATCAAAATTCAG 2264

482 GATTTTCAAGGAACACAATTTTCACTTTTCCCACTTTTCCCAATCATCAATCAAAATTCAG 541

2265 CATGCTGTTTCAAGAGAGTTCAGATACACCCGAGCTCGGATCAAAATTTTCAGGGCTTCA 2324

542 TATGCCGGGTTCAAGAGAGTTCAGATACACCCGAGCTCGGATCAAAATTTTCAGGGCTTCA 601

2325 TGGGAATTCATTTCACTCTG 2344

Db 602 TGGGAATTCATTTCAATCTG 621

RESULT 6

AI323253

LOCUS

DEFINITION

AI323253

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus (house mouse)

EST.

AI323253.1 GI:4057682

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

AI323253

```

Qy 1642 TTAACCTCTCAGGCCCCCTTCCCTTTTCATCAAGACAAATCTCTATCTTGTGCAACAAAT 1701
Db 61 TTAACCTCACCAGGCCCCCTTCCCTTTTCATCAAGACAAATCTCTCTATGCGACCAAT 120
Qy 1702 GGTCTTTGTCCTCTCTCAATGTCGTTTAAACCTGCTCAATTTGTCAACAGTACAAAAG 1761
Db 121 GGGCTCTGTCTCCCTCTCAATGTCGTTTAAACCTGCTCAATTTGTCAACAGTACAAAAG 180
Qy 1762 CAATTTAGGTATGAAGCCAGCTACAGATGTCAGATGTCAGGCTCCCTCAGATAATGAG 1821
Db 181 CAATTTAGGTACGAGATCAGCTCAGATGTCAGGCTCCCTCAGATAATGAG 240
Qy 1822 TACTTCTACGTGTATTTCAAGAAATATGAATATGATCTCAATGAGGATTTTCCAGAGAA 1881
Db 241 TACTTCTACGTGTATTTCAAGAAATATGAATATGATCTCAATGAGGATTTTCCAGAGAG 300
Qy 1882 AATTAGGTTTGGAGGTTACTAGGATCAGCTGCTTTTGAAGAAATGATGAACGACACA 1941
Db 301 AACTTAGGTTTGGAGGTTCTCGGGCTCTGGGCTTTCCGAGGATTTCCGAGAGAG 360
Qy 1942 GCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAGGTTGCGCTCAAAATGCTGAAAGAA 2001
Db 361 GCCTATGCTATTAGTAAACCGGAGTCTCAATCAGGTGGGTCGAGATCTTAAGAG 420
Qy 2002 AAACGACAGCTCTGAAGAGAGGCACTATGTCAGAACTCAAGATGATGACCCAGCTG 2061
Db 421 AAAGCTGACAGCTCTGAAGAAAGAGCTCTCATGTCGGAGCTCAAAATGATGACCCACTT 480
Qy 2062 GGAAGCCAGAGATATTTGAACCTGCTGGGGGCTGCACACTGTCAGGACCAATTTAC 2121
Db 481 GGACAATGATACCACTCGTGAGATTTGTTGGGGCTTCCACACTGCCAGGCGCCAGTGCAC 540
Qy 2122 TTGATTTTGAATACTGTTGTCATGTTGATCTTCAACTATCTAAGAAATGATAAGAGAA 2181
Db 541 TGGATTTTGAATATGTTGTCATGAGACCTCTCACTACTTAAGATTAAGAGAG 600
Qy 2182 AATTTTCAAGGACTTGGACAGAGATTTTCAAGAAACAAATTTCAAGTTTACCCCACT 2241
Db 601 AATGTTTCAAGGACTTGGACAGAGATTTTCAAGAAACAAATTTCAAGTTTACCTTACT 659
Qy 2242 TTCCAATCAGATCAAAATTTCCAGATGCTGCTGTTTCCAGAGAGT 2285
Db 660 TTCCAGCAGATCAAAATTTCCAGATGCTGCTGTTTCCAGAGAGT 703

```

```

RESULT 7
BF523018 604 bp mRNA linear EST 11-DEC-2000.
LOCUS UI-R-C2p-rg-f-10-0-UI-r1 UI-R-C2p Rattus norvegicus cDNA clone
DEFINITION UI-R-C2p-rg-f-10-0-UI 5', mRNA sequence.
ACCESSION BF523018
VERSION BF523018.1 GI:11631033
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine;
Rattus.
1 (bases 1 to 604)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL 9704477
MEDLINE 8889548
PUBMED
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEERF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
cDNA Library Preparation: M.B. Soares Lab Clone distribution:

```

clones will be available through Research Genetics ([www.resgen.com](http://www.resgen.com))  
 This clone is also available through the I.M.A.G.E. Consortium at  
 LNL (info@image.llnl.gov). IMAGE ID= 1793119  
 Seq primer: M13 Forward.

# FEATURES

```

source
Location/Qualifiers
1..604
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-C2p-rg-f-10-0-UI"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UI-R-C2p"
/note="vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-C2p
library is a subcloned library derived from the UI-R-C1
library, which is a subcloned library derived from the
UI-R-C0 library. The UI-R-C0 library consisted of a
mixture of individually tagged normalized libraries
constructed from rat placenta, adult lung, brain, liver,
kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day
embryo. The tag is a string of 3-5 nucleotides present
between the Not I site and the oligo-dT track which allows
identification of the library of origin of a clone within
the mixture. The subcloned library (UI-R-C2p) was
constructed as follows: PCR amplified cDNA inserts from
UI-R-C1 clones from which 3' ESTs had been derived was
used as a driver in a hybridization with the UI-R-C1
library in the form of single-stranded circles. The
remaining single-stranded circles (subcloned library) was
purified by hydroxyapatite column chromatography,
converted to double-stranded circles and electroporated
into DH10B bacteria (Life Technologies) to generate the
UI-R-C2p library. This procedure has been previously
described (Bonaldo, Lennon and Soares, Genome Research 6:
791-806, 1996)."
BASE COUNT 166 a 132 c 167 g 139 t
ORIGIN

```

```

Query Match 12.6%; Score 440; DB 10; Length 604;
Best Local Similarity 85.4%; Pred. No. 4.9e-46;
Matches 504; Conservative 0; Mismatches 80; Indels 6; Gaps 1;
Qy 2016 TGAAGAGAGGCACTCATGTCAAGAACTCAAGATGATGACCCAGCTGGAGCCACGAGAA 2075
Db 9 TGAAGAGAGGCGCTCATGGCTGAGCTCAAAATGATGACCCAGCTGGGCGCCATGACNA 68
Qy 2076 TATTGTGAACCTGCTGGGGCGTGACACTGTGACGACCAATTTTACTTTGTAATA 2135
Db 69 CATCGTGAACCTGCTGGGGCGATGACACTGTGAGGCGCAGTGTTGATTTTGAATA 128
Qy 2136 CTGTTGCTATGTTGATCTTCTCAACTATCTAAGAGTAAAGAGAAAAATTTCAAGAC 2195
Db 129 TTGTTGTCATGTTGACCTGCTCAACTCTAAGAGTAAAGAGAAAAATTTCAAGAC 188
Qy 2196 TTGACACAGATTTTCAAGGACACATTTTCAAGTTTACCCCACTTCCCAATCAGATCC 2255
Db 189 GTGACACAGATTTTAAAGGAAATATTTTCAAGTTTACCCCACTTCCCAATCAGATCC 248
Qy 2256 AAATTCAGCATGCTGTTTCAAGAGAAATTCAGATACACCCGAGCTCGGATCAATCTC 2315
Db 249 AACTTCAGTATGCGGGTTTCAAGAGAAATTCAGATATACCCGCGCTGGATCAGTTCTC 308
Qy 2316 AGGCTTCATGGGAAATTCATTTTCACTCTGAAGATGAAATTTGAATATGAAACCAAAAAAG 2375
Db 309 AGGCTTCAATGGGAAATTCATTTTCACTCTGAAGATGAAATTTGAATATGAAACCAAGAG 368
Qy 2376 GCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2429
Db 369 GCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 428
Qy 2430 TGCATATCAAGTTTCCCAAGGAATGGAATTTCTGGAATTTAAGTCGTGTGTTCAAGAGA 2489

```

[illegible]



KEYWORDS EST.  
 SOURCE Sus scrofa (pig)  
 ORGANISM Sus scrofa  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 REFERENCE 1 (bases 1 to 507)  
 AUTHORS Fahrrenkrug, S.C., Smith, T.P.L., Freking, B.A., Cho, J., White, J.,  
 Vallet, J., Wise, T., Rohrer, G.A., Petrea, G., Sultana, R., Quackenbush,  
 J., and Keeler, J.W.  
 TITLE Porcine gene discovery by normalized cDNA-library sequencing and  
 EST cluster assembly  
 JOURNAL Mamm. Genome 13 (8), 475-478 (2002)  
 MEDLINE 22213789  
 PUBMED 12226715  
 COMMENT Contact: Smith TPL  
 USDA, ARS, US Meat Animal Research Center  
 PO Box 166, Clay Center, NE 68933-0166, USA  
 Tel: 402 762 4366  
 Fax: 402 762 4390  
 Email: smith@email.marc.usda.gov  
 Single pass sequencing. Bases called and alt trimmed with phred  
 v0.980904.e. Vector identified by cross\_match with the -minscore 18  
 and -minmatch 12 options.  
 PCR Primers  
 FORWARD: AGGAACAGCTATGACCAT  
 BACKWARD: GTTTCACGATCAGCAGC  
 Plate: 2 row: D column: 23  
 Seq primer: ATTAGTGACACTATAG.  
 FEATURES  
 source Location/Qualifiers  
 1..507  
 /organism="Sus scrofa"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9823"  
 /tissue\_type="pooled"  
 /lab\_host="DH109"  
 /clone\_lib="MARC 2PFG"  
 /notes="Vector: PCMV SPORT6; Site 1: NotI; Site 2: SalI;  
 Library made from pooled tissue from testis, ovary,  
 endometrium, hypothalamus, pituitary, and placenta."  
 BASE COUNT 144 a 132 c 118 g 112 t 1 others  
 ORIGIN  
 Query Match 11.1%; Score 389.6; DB 12; Length 507;  
 Best Local Similarity 88.0%; Pred. No. 1.le-39;  
 Matches 447; Conservative 0; Mismatches 59; Indels 2; Gaps 2;  
 QY 1194 CAGATGACGTGGACCTTCTCTCGAAATCATTTCTGTGACCAAAAGGCTTGTATAA 1253  
 DB 1 CAGGTGACGTGGACCTTCTCTCGAAA-CATTTCTTGTGACCAAGGGGACTCGACGA 59  
 QY 1254 CGGATACAGCATATCCAAAGTTTTCGAATCATAAGCCAGCCAGGAGAAATATATATCCA 1313  
 DB 60 CGGGTACAGCATATCTAAGTTTTCGAACCAATAACACCAGCCAGGACATACATCTCCA 119  
 QY 1314 TGCA-GAAATGATGATGCCAATTTACCAAAATGTTTCAGCTGAATATAAGAGGAAC 1372  
 DB 120 TGCANGAAAAATGACGACGCCACCTTACAAAGATGTTACGCTGAATATAAGAGGAAC 179  
 QY 1373 CTCAGTGTCCGAGAGCATCGCAAGTCAGCGCTGCTGTTCTCGATGGATACCCAT 1432  
 DB 180 CGCAAGTGTGGCCGAGGATCGCAAGTCAGCGCTTCTCTCTGATGGTTACCCGT 239  
 QY 1433 TACCATCTGGACCTGGAAGAAGTGTTCAGACAAGTCTCCCAACTGCACAGAAGAGATCA 1492  
 DB 240 TACCGTCTTGGACCTGGAAGAAGTGTTCAGACAAGTCTCCCAACTGCACAGAAGATCA 299  
 QY 1493 CAGAAAGAGTCTGGAATAGAAAGGCTTAACAGAAAAGTGTTCGACAGTGGGTGTGAGCA 1552  
 DB 300 CGGAAGGCATCTGGAATAAAAGGCCCAACAGAGAAAGTATTTCGACAGTGGATTTCCAGCA 359  
 QY 1553 GTACTCTAAACATGAGTGAAGCCATAAAAGGGTCTCTGGTCAAGTGTGTCATACAAAT 1612  
 DB 360 GCACCTCTGAACATGAGCGAGGCTGTCAAAAGGGTTCCTGAGTGAAGTGTGTGCTCAAAAT 419

QY 1613 CCCTTGGCAGATCTGTGAGACGATCCTTTTAAACTCTTCAGGCCCCCTTCCCTTTTCATCC 1672  
 DB 420 CCCTGGGCAGCTGTGTGAACGATCTCTTAAACTCACCAGGCCCTTCCCTTTTCATCC 479  
 QY 1673 AAGACACATCTCTATCTATGCAACAAT 1700  
 DB 480 AGGACACATCTCTATCTATGCAACAAT 507  
 RESULT 12  
 CB220941  
 LOCUS CB220941  
 DEFINITION lAbo28C03 Bos taurus Abomasum #1 library Bos taurus cDNA similar to  
 FLT3 receptor tyrosine kinase - GO terms: protein tyrosine kinase  
 (0004713); receptor (0004872);, mRNA sequence.  
 ACCESSION CB220941  
 VERSION CB220941.1 GI:28291455  
 KEYWORDS EST.  
 SOURCE Bos taurus (cow)  
 ORGANISM Bos taurus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;  
 Bovidae; Bovinae; Bos.  
 REFERENCE 1 (bases 1 to 510)  
 AUTHORS Hansen, C., Fu, A., Meng, Y., Li, C., Okine, E., Sensen, C.W., Gordon,  
 P.M.K. and Moore, S.S.  
 TITLE Gene Expression Profiling of the Bovine Gastrointestinal Tract  
 JOURNAL Unpublished  
 COMMENT Contact: Dr. Stephen Moore  
 Beef Genomics Laboratory  
 Dept of AFNS, University of Alberta  
 410 Agri/Fox, Dept of AFNS, U of A, Edmonton, AB, T6G 2P5, Canada  
 Tel: 780 492 0169  
 Fax: 780 492 4265  
 Email: stephen.moore@ualberta.ca  
 Insert Length: 510 Std Error: 0.00  
 POLYA=No.  
 FEATURES  
 source Location/Qualifiers  
 1..510  
 /organism="Bos taurus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9913"  
 /tissue\_type="Smooth muscle"  
 /cell\_type="Epithelial"  
 /dev\_stage="Young adult"  
 /lab\_host="XLI-BlueWRF"-strain"  
 /clone\_lib="Bos taurus Abomasum #1 library"  
 /notes="Organ: Abomasum; Vector: Uni-2ZAPXR; Site 1: EcoR  
 I; Site 2: Xho I"  
 BASE COUNT 142 a 120 c 107 g 141 t  
 ORIGIN  
 Query Match 10.9%; Score 382; DB 14; Length 510;  
 Best Local Similarity 85.7%; Pred. No. 1e-38;  
 Matches 437; Conservative 0; Mismatches 70; Indels 3; Gaps 1;  
 QY 2644 GGCATCTACACCAATTAAGATGATGCTGTGTCATATGGAATATCTGTGGAAATCTTC 2703  
 DB 1 GGCATCTACACCAATTAAGATGATGCTGTGTCATATGGAATATCTGTGGAAATATTC 60  
 QY 2704 TCACCTTGGTGTGAATCCTTACCTGGCATTCGCGTTGATCTTAACCTTCTACAAACTGATT 2763  
 DB 61 TCCTTGGTGTGAATCCTTACCTGGCATTCGCGTTGATCTTAACCTTCTACAAACTCATC 120  
 QY 2764 CAAAATGGATTTAAATATGGATCAGCCATTTTATGTACAGAGAAATATACATTAATG 2823  
 DB 121 CAGAGTGGGTTCAGATGACGACGACCATTTATGCCACGAGAAATATATCTTTAATG 180  
 QY 2824 CAATCTGCTGGGCTTTTGACTCAAGGAAACGGCCATTCCTTCCCTAAATTTGACTTCGTTT 2883  
 DB 181 CAATCTGCTGGGCTTTTGACTCCAGGAAACGACCATTCCTTCCCTAACTGACTTCATTT 240

2884 TTAGGATGTCAGCTGGCAGATGCAGAGAGCGATGTATCAGATGTGGATGGCCGTGTT 2943  
241 TTAGGGTGTGTCAGCTGGCAGATGCAGAGAGCGATGTATCAGATGTGGATGGCCGTGTT 300  
2944 TCGGAATGTCTCTACACCTACCAAAACAGGCGACCTTTTCAGCAGAGAGATGGATTTGGGG 3003  
301 TCAGAGGTCCTCTGTTTACCAAAACAGGCGACCTTTACAGCAGAGAGATGGATTTGGGG 360  
3004 CTACTCTCTCCGAGGCTCAGGTCGAAATTCGTAGAGGAAACAAATTTAGTTTAAAGGACT 3063  
361 CTACCCCTCTCTCAGGTTTCAGAGTGAAGATTCCTTAGAGAAACGATTTAGTCTTAAGGACC 420  
3064 TCATCCCTCCACCTATCCCTAAACAGG---CTGTAGATACCAAAACAGATTAATTCAT 3120  
421 TCATCCCTCCCTGCTACCTTAACAGAGCTGTGTAGATTTACCAAAATAAGATTAATTCAT 480  
3121 CACTAAAAGAAATCTATTATCAACTGCTG 3150  
481 CACTAAGTGAATATATTCTCAACTGCTG 510

RESULT 13  
BX279530  
LOCUS  
DEFINITION BX279530 NCI CGAP Kid11 Homo sapiens cDNA clone IMAGE9998A067500 ;  
IMAGE:3033869, mRNA sequence.  
ACCESSION BX279530  
VERSION BX279530.1 GI:28612110  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 492)  
AUTHORS Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,  
Radelof, U., Schneider, D. and Korn, B.  
TITLE Human UnigeneSet - RZPD3  
JOURNAL Unpublished  
COMMENT Contact: Ina Rolfs  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany  
RZPD; IMAGE998A067500.  
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;  
Human UnigeneSet - RZPD3 (RZPDLIB NO.972)  
http://www.rzpd.de/CloneCards/cgi-  
bin/showlib.pl.cgi?response?libNo=972 Contact: Ina Rolfs  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
Heubnerweg 6, D-14059 Berlin, Germany  
Tel: +49 30 32639 101  
Fax: +49 30 32639 111  
www.rzpd.de  
This clone is available royalty-free from RZPD;  
contact RZPD (clone@rzpd.de) for further information. Seq primer:  
ML3r, Primer sequence: TTTCAACAGGAACAGCTATGAC.

FEATURES  
source  
1. .492  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clones="IMAGE998A067500 ; IMAGE:3033869"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Kid11"  
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with  
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;  
Plasmid DNA from the normalized library NCI CGAP Kid3 was  
prepared, and ss circles were made in vitro. Following HAP  
purification, this DNA was used as tracer in a subtractive  
hybridization reaction. The driver was PCR-amplified cDNAs  
from a pool of 5,000 clones made from the same library  
(cloneIDs 1322376-1323911, 1456007-1456775, and  
1500552-1502855). Subtraction by Bento Soares and M.  
Fatima Bonaldo. "

BASE COUNT 147 a 102 c 92 g 149 t 2 others

ORIGIN

Query Match 10.5%; Score 367.4; DB 13; Length 492;  
Best Local Similarity 99.2%; Pred. No. 6.9e-37;  
Matches 368; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2916 GATGTATCAGAAATGTGGATGGCCGTGTTTCGGAATGTCCCTACACCTACCAAAACAGGCG 2975  
DB 122 GATGTATCAGAAATGTGGATGGCCGTGTTTCGGAATGTCCCTACACCTACCAAAACAGGCG 181  
QY 2976 ACCTTTTCAGCAGAGAGATGGATTTGGGGCTACTCTCTCCGAGGCTCAGSTCGAAGATTC 3035  
DB 182 ACCTTTTCAGCAGAGAGATGGATTTGGGGCTACTCTCTCCGAGGCTCAGSTCGAAGATTC 241  
QY 3036 GTAGAGGAAACAATTTAGTTTAAAGGACTTCATCCCTCCACCTATCCCTAAACAGGCTGTAG 3095  
DB 242 GTAGAGGAAACAATTTAGTTTAAAGGACTTCATCCCTCCACCTATCCCTAAACAGGCTGTAG 301  
QY 3096 ATTACCAAAACAGATTAATTTTCATCACTAAAGAAATCTATTATCAACTGCTGCTTCA 3155  
DB 302 ATTACCAAAACAGATTAATTTTCATCACTAAAGAAATCTATTATCAACTGCTGCTTCA 361  
QY 3156 CCAGACTTTTCTAGAGCCGCTCTGCTTTACTCTTGTGTTTCAAAGGACTTTTGTAAA 3215  
DB 362 CCAGACTTTTCTAGAGCCGCTCTGCTTTACTCTTGTGTTTCAAAGGACTTTTGTAAA 421  
QY 3216 ATCAAAATCATCTGTCTACAAAGGAGGAGGAGCTGATTAATGAACCTTTTATGGAGCATTGAT 3275  
DB 422 ATCAAAATCATCTGTCTACAAAGGAGGAGGAGCTGATTAATGAACCTTTTATGGAGCATTGAT 481  
QY 3276 CTGCAATCCAAG 3286  
DB 482 CTGCAATCCNAG 492

RESULT 14  
BU696796  
LOCUS  
DEFINITION BU696796 500 bp mRNA linear EST 09-OCT-2002  
LL2in11124T7 Hematopoietic Stem Cell Subtracted Library Mus  
musculus cDNA 5' similar to human flk-2, mRNA sequence.  
ACCESSION BU696796  
VERSION BU696796.1 GI:23607570  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 500)  
AUTHORS Phillips, R.L., Ernst, R.E., Brunk, B.P., Ivanova, N., Mahan, M.A.,  
Deanehan, J.K., Moore, K.A., Overton, G.C. and Lemischka, I.R.  
TITLE The genetic program of hematopoietic stem cells  
JOURNAL Science 288 (5471), 1635-1640 (2000)  
MEDLINE 20295303  
PUBMED 10834841  
COMMENT Contact: Lemischka, Ihor R.  
Department of Molecular Biology  
Princeton University  
Lewis Thomas Laboratory, Washington Road, Princeton, NJ 08544, USA  
Tel: 609 258 2838  
Fax: 609 258 2759  
Email: ilemischka@molbio.princeton.edu  
These ESTs are derived from a subtracted cDNA library enriched for  
gene products expressed in day 14-14.5 fetal liver hematopoietic  
stem cells defined as Lineageneg/lo, A4.1pos, ckitpos, Ly6A/E  
(Sca-1)pos  
Seq primer: ML3Reverse or T7.  
Location/Qualifiers  
1. .500  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/tissue\_type="Hematopoietic"

FEATURES  
source  
1. .500  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/tissue\_type="Hematopoietic"



```
/cell_type="Stem Cells; Lineagenes/lo, AA4.lpos, ckitpos,
ly6a/E (Sca-1)pos"
/dev stage="Embryonic day 14-14.5"
/lab_host="DH10B"
/clone_lib="Hematopoietic Stem Cell Subtracted Library"
/notes=Organ: Fetal liver; Vector: Sport 1; Site: 1; Sal I;
Site 2: Not 1; Two directionally cloned cDNA libraries
were made from fetal liver hematopoietic stem cells
enriched to be Lineagenes/lo, AA4.lpos, ckitpos, ly6a/E
(Sca-1)pos called Scapos and from AA4.lneg fetal liver
cells. Subtractive hybridization was performed by
hybridization of the target, Scapos, single stranded cDNA
library in pSport1 to biotinylated RNA transcribed with
the driver, AA4.lneg cDNA library in pSport2 with inserts
cloned in the complementary orientation. For detailed
protocols and additional information please see our
website at http://stemcell.princeton.edu."
BASE COUNT 126 a 110 c 141 g 120 t 3 others
ORIGIN

Query Match 10.2%; Score 358.4; DB 13; Length 500;
Best Local Similarity 84.8%; Pred No. 9.2e-36;
Matches 401; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

OY 2380 GAAGAAGAGGAGGACTTGAATGCTTACATTTTGAAGATCTTCTTGTGCTTGCATATCAA 2439
|||||
Db 26 GAAGAGGAGGAGGAGGATTTGACGTGCTGACGTTTGAGNCCCTCTTGTGCTGCTACCAA 85
|||||

OY 2440 GTTGGCCAAAGGATGGAATTTCTGGAATTTAAAGTCGTGTGTACAGAGACCTGGCCGCC 2499
|||||
Db 86 GTGGCCAAAGGAGGATGGAATTTCTGGAATTTCAAGTCGTGTGTGTACAGAGACCTGGCCGCC 145
|||||

OY 2500 AGGAAGCTGCTTCTACCCACGGAAGTGGTGAAGATATGTGACTTTGGATTGGCTCGGA 2559
|||||
Db 146 AGGAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 205
|||||

OY 2560 GATATCATGATGATTCACCACTATGTTGTACGGGGCAATGCCGTCTGCTGTAAATGG 2619
|||||
Db 206 GACATCTGTAGCGATCTTCCAGTCTGCTGAGGGCAACGCGGTGCGGGTGAAGTGG 265
|||||

OY 2620 ATGGCCCCGAAAGCCCTGTTTGAAGGCATCTACACCATTAAGAGTGTGCTGTCTATAT 2679
|||||
Db 266 ATGGCACTTGTGAGGCTTATTTGAAGGATCTACACATCAAGAGTGTGCTGTCTCTAC 325
|||||

OY 2680 GGAATATTACTGTGGAATCTTCTCACTTGTGTGAATCCCTTACCTTGGCATTCGGTT 2739
|||||
Db 326 GGCATCTCTCTGTGGAGATATTTTCACTGGGTGTGAACCTTACCTTGGCATTCCTGTC 385
|||||

OY 2740 GATGCTTAATCTTACAACTGATTCAAAATGGAATTTAAATGATCAGCCATTTTATGCT 2799
|||||
Db 386 GACGCTAATCTTCTAATACTGATTCAGATGGAATTTAAATGAGGAGGATTCATGCTG 445
|||||

OY 2800 ACAGAAGAATATACATTAATTAATGCAATCTCTGCTGGCTTTTTCACATCAAGGAA 2852
|||||
Db 446 ACAGAAGGATATACATTTGTAATGCAATCTCTGCTGGCTTTTTCACATCAAGGAA 498
|||||

RESULT 15
AK080326
LOCUS
DEFINITION
AK080326 3616 bp mRNA linear HTC 05-DEC-2002
Mus musculus 3 days neonate thymus cDNA, RIKEN full-length enriched
library, clone:A630054K21 product:colony stimulating factor 1
receptor, full insert sequence.
ACCESSION
AK080326
VERSION
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
```

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

Meth. Enzymol. 303, 19-44 (1999)  
99279253  
10349636

2

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
11042159

3  
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kutsuna, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861

4  
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojohori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Giesi, C., King, B., Kochiwa, H., Kueh, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staudli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y. Functional annotation of a full-length mouse cDNA collection Nature 409 (6821), 685-690 (2001)  
21085660  
11217851

5  
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)  
6 (bases 1 to 3616)  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akaira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,

URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216

COMMENT  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MPC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.  
Please visit our web site for further details.  
URL: <http://genome.gsc.riken.go.jp/>  
URL: <http://fantom.gsc.riken.go.jp/>

# FEATURES

source

```
1. .3616
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:A630054K21"
/db_xref="taxon:10090"
/clone="A630054K21"
/tissue_type="thymus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="3 days neonate"
76. .3014
/notes="colony stimulating factor 1 receptor
(MGD|MG1:1339758, GB|NM_007779, evidence: BLASTN, 98%,
match=3661)
putative"
```

BASE COUNT 854 a 998 c .972 g 792 t

## ORIGIN

```
Query Match 9.7%; Score 341.2; DB 11; Length 3616;
Best Local Similarity 58.5%; Pred. No. 4.8e-34;
Matches 674; Conservative 0; Mismatches 463; Indels 15; Gaps 4;

QY 1715 TCTTCATTGTCGTTTAAACCTGCTAATTTGTCAACAGTACAAAAAGCAATTTAGGTATG 1774
DB 1648 TCATGTTTCTGCTGGTGCTACTGCTGTGCTCTGTGTACAGTACAGCAGAGCCGA 1707

QY 1775 AAAGCAGCTACAGATGTTACAGGTGACCGGCTCTCAGATATAGTACTTCTAGCTTG 1834
DB 1708 AGTACCAGGTGCGCTGGAAGTATCATCGAGATACGAAGGCAATAGTACACCTTCATTG 1767

QY 1835 ATTTCAGAGATATGATATGATCTCAATGGGAGTTTCCAAAGAGAAATTTAGAGTTTG 1894
DB 1768 ACCCTACTCAGTTGCCCTCAATGAGAGTGGGAGTTCCCTCGGAACAACCTGCAGTTTG 1827

QY 1895 GGAAGGTACTAGGATCAGGTGCTTTTGGAAAAGTGATGAACGCAACAGCTTATGGAATTA 1954
DB 1828 GTAAGACTTAGAGCGCGTGCTTTGGNAGGTGGTGGAGGCTACAGCCTTTGGTCTGG 1887

QY 1955 GCAAAAACAGAGTCTCAATCCAGGTTGCCGTCAAAATGCTGAAAGAAAAAGCAGACGCT 2014
DB 1888 GCAAGAAGATGCAAGTGGCTGAAAGTGGCTGTGAAGATGCTAAAGTCCACGGCTCATGCTG 1947

QY 2015 CTGAAGAGAGGCACTCATGTGCAAGTCAAGATGATGATGATGATGATGATGATGATGAT 2074
DB 1948 ATGAGAAGAGAGGCGCTGATGTGAGAGTGAAGATCATGAGTCACTTGGGAGCAGCAGAGA 2007

QY 2075 ATATTGTGAACCTGCTGGGGCGGTGACACTGTGAGGACCAATTTTACTTCAATTTTGAAT 2134
DB 2008 ATATAGTCAACCTTTGGGAGCGCTGTACTCAGGAGACCTGTCTCTGGTCACTCACTGAAT 2067

QY 2135 ACTGTTGCTATGCTGATCTTCTCAACTATCTAAAGAGTAAAGAGAAAAATTTTACAGGA 2194
DB 2068 ACTGCTGCTATGAGACCTACTCAACTTCTCCGAG- GAAGGCCGAGGCTATGCTAGGA 2126

QY 2195 CTGGAACAGATTTTCAAGGAACAATTTTCAAGTTTATCCCACTTTTCCCAATCACATC 2254
DB 2127 CCAGCCTGAGTCTGCTGAGGACTCCGAGGGAGACTCCAGC-----TACAAGAACATC 2180
```

Search completed: August 28, 2003, 05:42:18  
Job time : 8457.93 secs

```
QY 2255 CAAATTCACGATGCTCGTGGTTCAAGAGAGAGTTTCAACCCGAGACTCGGATCAAAATCT 2314
DB 2181 CACCTGAGAGAAATATATGTGCGCAGGACAGATGGCTTCTCCAGTCAAGGTTGAGACACC 2240

QY 2315 CAGGGCTTATGGGAATTCATTTCACTCTCAAGATGAATGAATGAATGAATGAATGAATGA 2374
DB 2241 TAGCTGGAGATGAGGCGCTGTCTGACT-----TCTTCAAGTGAATCTCTTTTAAGCAAG 2295

QY 2375 GGCTGGAAGAAGAGAGGAGGACTTGAATGCTGCTTACATTTGAAGATCTCTTTTGGCTTGCAT 2434
DB 2296 ATCTGACAAGAGAGGCGCAGCGCG- ---CCCTGGAGCTCTGGGACCTGCTCACTTCTCCA 2352

QY 2435 ATCAAGTTCGCAAGAGGAATTTCTGGAATTTAAGTCTGTGTGTTTCAAGAGACCTGG 2494
DB 2353 GCCAAGTGGCTCAGGGCATGGCCTTCTCTTCTTAAAACTGCATCCACCGGACGTAG 2412

QY 2495 CCGCCAGGAACGCTGCTGTCTACCCACCGGAAAGTGGTGAAGATATGTGACTTTGGATTGG 2554
DB 2413 CAGCTCGAAACGCTGCTGTTTGACCAAGGAGCATGTGGCCAAAGATTGGGAGCTTTGGACTGG 2472

QY 2555 CTGAGATATCATGAGTGATTTCCAACTATGTGTGTCAGGGGCAATGCCGCTCTGCTGTAA 2614
DB 2473 CTAGGGACATCATGATGACTCCAACTATGTGTTCAGGGGCAATGCCGCTCTGCTGTAA 2532

QY 2615 AATGGATGGCCCCCGAAAGCCTGTTTGAAGGCATCTACACCATTTAAGAGTGAATGCTGGT 2674
DB 2533 AGTGGATGGCCCCCAGAGAGCATCTTTGACTGCGTCTACACAGTTTACAGATGATGTGGT 2592

QY 2675 CATATGGAATATTACTGTGGGAAATCTTCTCACTTGTGTGTAATCCTTACCTGGCATTC 2734
DB 2593 CCTACGGCATCTCTCTCTGGGAGATCTTCTCGTGTGTTGTTCTGAACCCCTACCCCGCATCC 2652

QY 2735 CGTTTGATGCTAACTTCTACAAACTGATTTCAAAATGGATTTAAATTCGATCAGCCATTTT 2794
DB 2653 TAGTGAACAAAGTTTCTCAAAACTGGTGAAGATGATACCAATGGCCAGCCTGTAT 2712

QY 2795 ATGCTACAGAAGAAATATATACATTTAATGCAATCTCTCTGGGCTTTTGAATCAAGGAAAC 2854
DB 2713 TTGCACCGAAGACATATACAGCATCATGAGTCTCTGCTGGGACCTGGAGCCTACAGAA 2772

QY 2855 GGCCATCTCTTC 2866
DB 2773 GACCCACCTTTC 2784
```

[illegible]

PR 18-JUN-1993; 93US-0080244.  
 PR 21-JUN-1993; 93US-0081508.  
 PR 23-NOV-1993; 93US-0157490.  
 PA (UYP-.) UNIV PRINCETON.  
 XX Lemischka IR;  
 PI WPI; 1995-052014/07.  
 DR P-PSDB; AAR67816.  
 XX  
 PT Ligand for receptor protein tyrosine kinase - useful for the  
 PT stimulation of primitive haematopoietic stem cells causing  
 PT proliferation and/or differentiation  
 XX  
 PS Disclosure; Fig 1b; 131pp; English.  
 XX  
 CC The sequence corresponds to a cDNA encoding a human Flk2 (fetal  
 CC liver kinase) receptor protein-tyrosine-kinase. Flk2 is expressed  
 CC in primitive hematopoietic cells but not in mature hematopoietic  
 CC cells. The gene product is useful in isolation of receptor  
 CC ligands, which have applications in diagnosis of bone marrow  
 CC disorders and in stimulating proliferation and/or differentiation  
 CC of primitive hematopoietic stem cells.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 3501 BP; 1068 A; 709 C; 784 G; 940 T; 0 other;  
 Query Match 100.0%; Score 3501; DB 16; Length 3501;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGAGGGCGCATCCGAGGGCTGGGCGCGCCCTCGGGGACCCCGGGCTCCGGAGGCCATG 60  
 DB 1 CGAGGGCGCATCCGAGGGCTGGGCGCGCCCTCGGGGACCCCGGGCTCCGGAGGCCATG 60  
 QY 61 CCGCGCTGGCGCGCGACGGGGCACCGTGGCTGCTGCTGTTTTTCTGCAATGATA 120  
 DB 61 CCGCGCTGGCGCGCGACGGGGCACCGTGGCTGCTGCTGTTTTTCTGCAATGATA 120  
 QY 121 TTTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTTTAATCATATAAG 180  
 DB 121 TTTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTTTAATCATATAAG 180  
 QY 181 AACATGATTCATCAGTGGGAGTCAATCATATCCATGGTATCAGAAATCCCGGAA 240  
 DB 181 AACATGATTCATCAGTGGGAGTCAATCATATCCATGGTATCAGAAATCCCGGAA 240  
 QY 241 GACCTCGGGTGTGGTTGAGACCCCGAGAGCTCAGGGACAGTGTACGAAGCTGCCGCTGTG 300  
 DB 241 GACCTCGGGTGTGGTTGAGACCCCGAGAGCTCAGGGACAGTGTACGAAGCTGCCGCTGTG 300  
 QY 301 GAAGTGAATGATCTGCTTCCATCAGCTGCAAGTGTGCTGATGCCCCAGGGAAATTT 360  
 DB 301 GAAGTGAATGATCTGCTTCCATCAGCTGCAAGTGTGCTGATGCCCCAGGGAAATTT 360  
 QY 361 TCCTGTCTGGGCTTTAAGCACAGCTCCCTGAAATGCGAGCCACATTTTGAATTTACAA 420  
 DB 361 TCCTGTCTGGGCTTTAAGCACAGCTCCCTGAAATGCGAGCCACATTTTGAATTTACAA 420  
 QY 421 AACAGAGGAGTTGTTTCCATGGTCAATTTGAAATGACAGAAACCCAGCTGGAGATAC 480  
 DB 421 AACAGAGGAGTTGTTTCCATGGTCAATTTGAAATGACAGAAACCCAGCTGGAGATAC 480  
 QY 481 CTACTTTTATTACAGTGAAGCTACCAATTTACCAATTTGTTTACAGTGAATTAAGA 540  
 DB 481 CTACTTTTATTACAGTGAAGCTACCAATTTACCAATTTGTTTACAGTGAATTAAGA 540  
 QY 541 AATACCTGCTTTTACATTAAGAGACCTTACTTTAGAAAAATGAAAAACGAGAGCC 600  
 DB 541 AATACCTGCTTTTACATTAAGAGACCTTACTTTAGAAAAATGAAAAACGAGAGCC 600  
 QY 601 CTGGTCTGCAATCTGAGAGCGTTCCAGAGCCGATCGTGGAAATGGTGTCTTGGGATTCA 660

DB 601 CTGGTCTGCAATCTGAGAGCGTTCCAGAGCCGATCGTGAAATGGGTGCTTGGGATTC 660  
 QY 661 CAGGGGAAAGCTGTTAAAGAAAGAGTCCAGCTGTTGTTTAAAGAGGAGGAAAAAGTGTCT 720  
 DB 661 CAGGGGAAAGCTGTTAAAGAAAGAGTCCAGCTGTTGTTTAAAGAGGAGGAAAAAGTGTCT 720  
 QY 721 CATGAATTTATTTGGGACGACATAAAGTGTGTTGCCAGAAATGAACCTGGGAGGAAATGC 780  
 DB 721 CATGAATTTATTTGGGACGACATAAAGTGTGTTGCCAGAAATGAACCTGGGAGGAAATGC 780  
 QY 781 ACAGGCTGTTCAATAGATCTAAATCAAACTCCTCAGACCACATTTGCCCAATTTATTT 840  
 DB 781 ACCAGGCTGTTCAATAGATCTAAATCAAACTCCTCAGACCACATTTGCCCAATTTATTT 840  
 QY 841 CTTAAAGTAGGGAAACCCCTTATGGATAAGGTGCAAGTGTGTTGCCAGAAATGAACCTGGGAGGAAATGC 900  
 DB 841 CTTAAAGTAGGGAAACCCCTTATGGATAAGGTGCAAGTGTGTTGCCAGAAATGAACCTGGGAGTTC 900  
 QY 901 GGGCTCACCTGGGAATTTAGAAAAACAAAGCACTCGAGGAGGCACTACTTTGAGATGAGT 960  
 DB 901 GGGCTCACCTGGGAATTTAGAAAAACAAAGCACTCGAGGAGGCACTACTTTGAGATGAGT 960  
 QY 961 ACCTATTTCAACAAACAGAACTATGATACGGAATCTGTTTGTGTTTGTATCATCAGTGCA 1020  
 DB 961 ACCTATTTCAACAAACAGAACTATGATACGGAATCTGTTTGTGTTTGTATCATCAGTGCA 1020  
 QY 1021 AGAAACGACACCGGATACACTACTGTTTCTCTTCAAGACATCCAGTCAATCAGCTTTG 1080  
 DB 1021 AGAAACGACACCGGATACACTACTGTTTCTCTTCAAGACATCCAGTCAATCAGCTTTG 1080  
 QY 1081 GTTACCATCGTAGGAAAGGATTTATAAATGCTACCAATTCAGTGAAGATTTATGAAATTT 1140  
 DB 1081 GTTACCATCGTAGGAAAGGATTTATAAATGCTACCAATTCAGTGAAGATTTATGAAATTT 1140  
 QY 1141 GACCAATATGAGAGTGTGTTTTTCTGTCAGTGTGTTTAAAGCTTACCAACAAATCAGATGT 1200  
 DB 1141 GACCAATATGAGAGTGTGTTTTTCTGTCAGTGTGTTTAAAGCTTACCAACAAATCAGATGT 1200  
 QY 1201 ACCTGGAGCTTCTCTGAAATCAATTTCTGTCAGCAAAAGGCTCTGTAAACGGATAC 1260  
 DB 1201 ACCTGGAGCTTCTCTGAAATCAATTTCTGTCAGCAAAAGGCTCTGTAAACGGATAC 1260  
 QY 1261 AGCATATCCAAGTTTTCGAATCAATGACACAGGACGAGAAATATATATTCATGCGAGAA 1320  
 DB 1261 AGCATATCCAAGTTTTCGAATCAATGACACAGGACGAGAAATATATATTCATGCGAGAA 1320  
 QY 1321 AATGATGATGCCAAATTTTACAAATGTTTCACTGCTGAATATAAGAGGAAACCTCAAGTG 1380  
 DB 1321 AATGATGATGCCAAATTTTACAAATGTTTCACTGCTGAATATAAGAGGAAACCTCAAGTG 1380  
 QY 1381 CTGCGAGAGCATCGGCAAGTCTGAGCAAGTCTCCCACTGCAAGAGAGATACCAAGATACCT 1440  
 DB 1381 CTGCGAGAGCATCGGCAAGTCTGAGCAAGTCTCCCACTGCAAGAGAGATACCAAGATACCT 1440  
 QY 1441 TGGAACCTGGAAGAGTGTTCAGCAAGTCTCCCACTGCAAGAGAGATACCAAGAGAG 1500  
 DB 1441 TGGAACCTGGAAGAGTGTTCAGCAAGTCTCCCACTGCAAGAGAGATACCAAGAGAG 1500  
 QY 1501 GTCTGGAATGAAAGGCTTAAAGAGGTTTGGACAGTGGGTGTGAGCAGTACTCTA 1560  
 DB 1501 GTCTGGAATGAAAGGCTTAAAGAGGTTTGGACAGTGGGTGTGAGCAGTACTCTA 1560  
 QY 1561 AACATGATGAGGCAATAAAGGTTTCCGTCAGTGTGTCGATACAAATTCCTTGGC 1620  
 DB 1561 AACATGATGAGGCAATAAAGGTTTCCGTCAGTGTGTCGATACAAATTCCTTGGC 1620  
 QY 1621 ACATCTTGTGAGAGTCTTTTAAACTCTCCAGGCCCCCTTCCCTTTCATCCAGAGAAC 1680  
 DB 1621 ACATCTTGTGAGAGTCTTTTAAACTCTCCAGGCCCCCTTCCCTTTCATCCAGAGAAC 1680  
 QY 1681 ATCTCATTTCTATGCAACAAATGGTGTGTTGTCCTCTTCTTCTGTTTAAACCTGCTA 1740

Db 1681 ATCTATTCTATCAACAATTGGTGTGTGTCCTCTCTTCAATGTCGTTTTTAACCCCTGCTA 1740  
Qy 1741 ATTTGTACACAGTACAAAAAGCAATTTAGTGATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
Db 1741 ATTTGTACACAGTACAAAAAGCAATTTAGTGATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
Qy 1801 ACCGGCTCTCAGATAATAGTACTTCTACGTTGATTTTCAGAGAAATATGAATATGATCTC 1860  
Db 1801 ACCGGCTCTCAGATAATAGTACTTCTACGTTGATTTTCAGAGAAATATGAATATGATCTC 1860  
Qy 1861 AAATGGGAGTTTCCAAAGAGAAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTTT 1920  
Db 1861 AAATGGGAGTTTCCAAAGAGAAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTTT 1920  
Qy 1921 GGAAGAGTGAATGAAACGACAGCTTATGGAATTTAGCAAAACAGAGGTCTCAATCCAGGTT 1980  
Db 1921 GGAAGAGTGAATGAAACGACAGCTTATGGAATTTAGCAAAACAGAGGTCTCAATCCAGGTT 1980  
Qy 1981 GCGCTCAAAATGCTGAAGAAAGACAGACAGCTCTGAAGAGAGGCACTCATGTCCAGAA 2040  
Db 1981 GCGCTCAAAATGCTGAAGAAAGACAGACAGCTCTGAAGAGAGGCACTCATGTCCAGAA 2040  
Qy 2041 CTCAAGATGATGACCCAGCTGGGAAGCCACGAGATATTGTGAACCTGTGGGGGCTGC 2100  
Db 2041 CTCAAGATGATGACCCAGCTGGGAAGCCACGAGATATTGTGAACCTGTGGGGGCTGC 2100  
Qy 2101 ACACTGTCCAGGACCAATTTACTGATTTTGAATCTGTTGATGCTGATCTTCTCAAC 2160  
Db 2101 ACACTGTCCAGGACCAATTTACTGATTTTGAATCTGTTGATGCTGATCTTCTCAAC 2160  
Qy 2161 TATCTAAGAGTAAAGAGAAAAATTTACAGGACTTTGGACAGAGATTTTCAAGGAACAC 2220  
Db 2161 TATCTAAGAGTAAAGAGAAAAATTTACAGGACTTTGGACAGAGATTTTCAAGGAACAC 2220  
Qy 2221 TATTTTCAGTTTACCCCACTTTCCAAATCAGATCCAAATTTCCAGCATGCTGCTCAAGA 2280  
Db 2221 TATTTTCAGTTTACCCCACTTTCCAAATCAGATCCAAATTTCCAGCATGCTGCTCAAGA 2280  
Qy 2281 GAAGTTTCAGATACACCCGAGCTCGGATCAAATCTCAGGGCTTCATGGGAATTCATTTTCA 2340  
Db 2281 GAAGTTTCAGATACACCCGAGCTCGGATCAAATCTCAGGGCTTCATGGGAATTCATTTTCA 2340  
Qy 2341 TCTGAAGATGAATTTGAATATGAAACCAAAAAAGGCTGGAAGAGAGGAGGACTTTGAAT 2400  
Db 2341 TCTGAAGATGAATTTGAATATGAAACCAAAAAAGGCTGGAAGAGAGGAGGACTTTGAAT 2400  
Qy 2401 GTGCTTACATTTGAAGATCTTCTTTGCTTTGATATCAAGTTGCCAAAGGAATGGAATTT 2460  
Db 2401 GTGCTTACATTTGAAGATCTTCTTTGCTTTGATATCAAGTTGCCAAAGGAATGGAATTT 2460  
Qy 2461 CTGGAATTTAAGTCTGTGTTCACAGAGACCTGGCCGCCAGGAACGTTGTCACCCAC 2520  
Db 2461 CTGGAATTTAAGTCTGTGTTCACAGAGACCTGGCCGCCAGGAACGTTGTCACCCAC 2520  
Qy 2521 GGGAAAGTGGTGAAGATATGACCTTTGGAATGGCTCGAGATATCATGATGATTTCCAAAC 2580  
Db 2521 GGGAAAGTGGTGAAGATATGACCTTTGGAATGGCTCGAGATATCATGATGATTTCCAAAC 2580  
Qy 2581 TATGTTGTCCAGGGCAATCCCGCTGCTGTGAATATGATGATGATGATGATGATGATGATG 2640  
Db 2581 TATGTTGTCCAGGGCAATCCCGCTGCTGTGAATATGATGATGATGATGATGATGATGATG 2640  
Qy 2641 GAAAGCATCTACACCAATTAAGAGTATGATGCTGTATGATGATGATGATGATGATGATGATG 2700  
Db 2641 GAAAGCATCTACACCAATTAAGAGTATGATGCTGTATGATGATGATGATGATGATGATGATG 2700  
Qy 2701 TTCTCACTTGGTGTGAATCTTACCTTGGCAATTCGCGTTGATGCTAACTTTACAAACTG 2760  
Db 2701 TTCTCACTTGGTGTGAATCTTACCTTGGCAATTCGCGTTGATGCTAACTTTACAAACTG 2760  
Qy 2761 ATTCAAAATGGATTTAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2820  
Db 2761 ATTCAAAATGGATTTAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2820

Qy 2821 ATGCAATCTGCTGGGCTTTTGACTCAAGGAAACGGCCATCCTTCCCTTAATTTGACTTCG 2880  
Db 2821 ATGCAATCTGCTGGGCTTTTGACTCAAGGAAACGGCCATCCTTCCCTTAATTTGACTTCG 2880  
Qy 2881 TTTTGTAGATGTCAAGCTGGCAGATGCAAGAGAGCGATGATCAGAAATGTCGATGCGCCGT 2940  
Db 2881 TTTTGTAGATGTCAAGCTGGCAGATGCAAGAGAGCGATGATCAGAAATGTCGATGCGCCGT 2940  
Qy 2941 GTTTCGGAATGTCTCTCACTACCAAAACAGGCGACCTTTTCAGCAGAGAGATGATTTG 3000  
Db 2941 GTTTCGGAATGTCTCTCACTACCAAAACAGGCGACCTTTTCAGCAGAGAGATGATTTG 3000  
Qy 3001 GGGCTACTCTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGGAACAATTTAGTTTTAAGG 3060  
Db 3001 GGGCTACTCTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGGAACAATTTAGTTTTAAGG 3060  
Qy 3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTTAATTCAT 3120  
Db 3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTTAATTCAT 3120  
Qy 3121 CACTAAAAGAAAAATCTATTATCAACTGCTGCTTACCAGACTTTTCTCTAGAAGCCGTCT 3180  
Db 3121 CACTAAAAGAAAAATCTATTATCAACTGCTGCTTACCAGACTTTTCTCTAGAAGCCGTCT 3180  
Qy 3181 GCGTTTACTCTGTTTTTCAAAGGGACTTTTGTAAATCAAATCATCTCTGTCACAAGGCAG 3240  
Db 3181 GCGTTTACTCTGTTTTTCAAAGGGACTTTTGTAAATCAAATCATCTCTGTCACAAGGCAG 3240  
Qy 3241 GAGGAGCTGATATGAATCTTTATGGAGCATTTGATCTGCATCCAGGCCCTTCTCAGGCCG 3300  
Db 3241 GAGGAGCTGATATGAATCTTTATGGAGCATTTGATCTGCATCCAGGCCCTTCTCAGGCCG 3300  
Qy 3301 GCTTGAGTGAATGTCTGCTGAGTACAGTATATCTTCTGTAATACATAAAACAAAGC 3360  
Db 3301 GCTTGAGTGAATGTCTGCTGAGTACAGTATATCTTCTGTAATACATAAAACAAAGC 3360  
Qy 3361 ATTTTCTAAGGAGAAAGCTAAATATGATTTTTTAAAGTCTATGTTTTTAAATAATATGTAAA 3420  
Db 3361 ATTTTCTAAGGAGAAAGCTAAATATGATTTTTTAAAGTCTATGTTTTTAAATAATATGTAAA 3420  
Qy 3421 TTTTTCAGCTATTTAGTGATATATTTTATGGTGGGAATAAAAATTTCTACTACAGAAAAA 3480  
Db 3421 TTTTTCAGCTATTTAGTGATATATTTTATGGTGGGAATAAAAATTTCTACTACAGAAAAA 3480  
Qy 3481 AAAAAAAAAAAAAAAAAAAAAA 3501  
Db 3481 AAAAAAAAAAAAAAAAAAAAAA 3501

## RESULT 2

AAQ79069

ID AAQ79069 standard; cDNA; 3501 BP.

XX AAQ79069;

XX AC

XX 25-MAR-2003 (updated)

DT 04-JUL-1995 (first entry)

XX Human flk-2 cDNA.

DE

XX Fetal liver kinase-2; flk-2; protein tyrosine-kinase receptor;

KW hematopoiesis; stem cell; ds.

XX Homo sapiens.

OS

XX Key

FH CDS

FT /tag= a

FT sig\_peptide

FT 58..138

FT /tag= b

FT mat\_peptide

FT 139..3036

FT /tag= c

Location/Qualifiers

58..3039

/tag= a

58..138

/tag= b

139..3036

/tag= c







KW differentiation; treatment; anaemia; bone marrow damage;  
 XX cancer chemotherapy; radiation; ds.  
 OS Mus musculus.

XX PH Key Location/Qualifiers  
 XX CDS 58..3039  
 XX FT sig\_peptide /tag= a  
 XX FT mat\_peptide /tag= b  
 XX FT 139..3036  
 XX FT /tag= c

PN US5548065-A.

PD 20-AUG-1996.

XX 31-OCT-1994; 94US-0252517.

XX 19-NOV-1992; 92US-0977451.

XX 02-APR-1991; 91US-0679666.

XX 28-JUN-1991; 91US-0728913.

XX 15-NOV-1991; 91US-0793065.

XX 24-DEC-1991; 91US-0813593.

XX 26-JUN-1992; 92US-0906397.

XX 12-NOV-1992; 92US-0975049.

XX 30-APR-1993; 93US-0055269.

XX 31-OCT-1994; 94US-0252517.

PA (UYPR-) UNIV PRINCETON.

XX Lemischka IR;

XX WPI; 1996-392678/39.

XX P-ESDB; AAR97419.

XX Anti-fœtal liver kinase 2 (flk-2) antibodies - useful in assays,  
 XX for isolating haematopoietic stem cells expressing receptor and for  
 XX obtaining ligands  
 XX Disclosure; Columns 39-48; 50pp; English.

XX The present sequence encodes murine foetal liver kinase 2 (flk-2),  
 XX a protein tyrosine kinase. Isolated antibodies, pref. monoclonal,  
 XX raised against the extracellular portion of flk-2 can be used to  
 XX assay for flk receptors on the surface of primitive haematopoietic  
 XX stem cells, and to isolate positive cells. The antibodies can also  
 XX be used as, or to obtain ligands, which stimulate the proliferation  
 XX and/or differentiation of stem cells. The ligands can be used, e.g.  
 XX for treating anaemia, or bone marrow damage resulting from cancer  
 XX chemotherapy, or radiation.

XX (Updated on 25-MAR-2003 to correct PF field.)

XX SQ Sequence 3501 BP; 1068 A; 709 C; 784 G; 940 T; 0 other;

Query Match 100.0%; Score 3501; DB 17; Length 3501;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGAGCGGCATCCGAGGCTGGCGCGCCCTGGGGACCCCGGGCTCCGAGGCCATG 60

Db 1 CGAGCGGCATCCGAGGCTGGCGCGCCCTGGGGACCCCGGGCTCCGAGGCCATG 60

Qy 61 CCGCGTTGGCGCGGACCGGACCGTGGCGTGGTGTGTTTTCGCAATGATA 120

Db 61 CCGCGTTGGCGCGGACCGGACCGTGGCGTGGTGTGTTTTCGCAATGATA 120

Qy 121 TTTGGGACTATTCAAAATCAAGATCTGCTGATCAAGTGTGTTTTCGCAATGATA 180

Db 121 TTTGGGACTATTCAAAATCAAGATCTGCTGATCAAGTGTGTTTTCGCAATGATA 180

Qy 181 AACATGATTCATCAGTGGGGAAGTCATCATCATATCCCATGATCAGAAATCCCGGAA 240

Db 181 AACATGATTCATCAGTGGGGAAGTCATCATCATATCCCATGATCAGAAATCCCGGAA 240

Db 181 AACATGATTCATCAGTGGGGAAGTCATCATCATATCCCATGATCAGAAATCCCGGAA 240  
 Qy 241 GACCTCGGGTGTGGTTCAGAGCCCGAGAGCTCAGGACAGAGTGTACAGAGCTCCGCTGTG 300  
 Db 241 GACCTCGGGTGTGGTTCAGAGCCCGAGAGCTCAGGACAGAGTGTACAGAGCTCCGCTGTG 300  
 Qy 301 GAAGTGGATGATCTGCTTCCATCACTGCAAGTGTGGTGGATGCGCCAGGGAACATT 360  
 Db 301 GAAGTGGATGATCTGCTTCCATCACTGCAAGTGTGGTGGATGCGCCAGGGAACATT 360  
 Qy 361 TCCTGTCTCTGGGTCTTTAAGCAGAGCTCCCTGAATGCCAGCCACATTTGATTACAA 420  
 Db 361 TCCTGTCTCTGGGTCTTTAAGCAGAGCTCCCTGAATGCCAGCCACATTTGATTACAA 420  
 Qy 421 AACAGAGAGTGTGTTTCCATGTCATTTTGAATAATGACAGAAACCAAGCTCGAGAAATAC 480  
 Db 421 AACAGAGAGTGTGTTTCCATGTCATTTTGAATAATGACAGAAACCAAGCTCGAGAAATAC 480  
 Qy 481 CTACTTTTATTCAGAGTGAAGCTACCAATTAACAATATTGTTTACAGTGAATTAAGA 540  
 Db 481 CTACTTTTATTCAGAGTGAAGCTACCAATTAACAATATTGTTTACAGTGAATTAAGA 540  
 Qy 541 AATACCCCTGCTTTACACATTAAGAAGCTTTTACATTTAGAAAATGGAACACGAGCGCC 600  
 Db 541 AATACCCCTGCTTTACACATTAAGAAGCTTTTACATTTAGAAAATGGAACACGAGCGCC 600  
 Qy 601 CTGGTCTGCATATCTGAGAGCGTTCAGAGCGGATCGTGGAAATGCTGTTTCCGATTCA 660  
 Db 601 CTGGTCTGCATATCTGAGAGCGTTCAGAGCGGATCGTGGAAATGCTGTTTCCGATTCA 660  
 Qy 661 CAGGGGAAAGCTGTAAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAAGTCTT 720  
 Db 661 CAGGGGAAAGCTGTAAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAAGTCTT 720  
 Qy 721 CATGAATTTTGGGACGACATAAGTGTGTCGACAGAAATGAATGGGACAGGAATGC 780  
 Db 721 CATGAATTTTGGGACGACATAAGTGTGTCGACAGAAATGAATGGGACAGGAATGC 780  
 Qy 781 ACCAGGCTGTTCAACAATAGATCTAAATCAAACTCTCAGACACACATTGCCAAATTTT 840  
 Db 781 ACCAGGCTGTTCAACAATAGATCTAAATCAAACTCTCAGACACACATTGCCAAATTTT 840  
 Qy 841 CTTAAAGTAGGGAAACCCCTTATGGATAAGTGCAGAAAGCTGTTTCATGTGAACCATGATTC 900  
 Db 841 CTTAAAGTAGGGAAACCCCTTATGGATAAGTGCAGAAAGCTGTTTCATGTGAACCATGATTC 900  
 Qy 901 GGGCTCACTCGGAAATTAAGAAAACAAAGCACTCGAGGAGGGCAACTCTTTGAGATGAGT 960  
 Db 901 GGGCTCACTCGGAAATTAAGAAAACAAAGCACTCGAGGAGGGCAACTCTTTGAGATGAGT 960  
 Qy 961 ACCTATTCAACAAACAGAACTATGATACGGATCTGTTTGTGTTTGTATCATCATGCGCA 1020  
 Db 961 ACCTATTCAACAAACAGAACTATGATACGGATCTGTTTGTGTTTGTATCATCATGCGCA 1020  
 Qy 1021 AGAAACGACACCGGATCTACACTTGTTCCTCTTCAAGAGCATCCAGTCAATCAGCTTTG 1080  
 Db 1021 AGAAACGACACCGGATCTACACTTGTTCCTCTTCAAGAGCATCCAGTCAATCAGCTTTG 1080  
 Qy 1081 GTTACCATCGTAGGAAAGGATTTATAAATGCTACCAATTCAGTGAAGATTTATGAAATT 1140  
 Db 1081 GTTACCATCGTAGGAAAGGATTTATAAATGCTACCAATTCAGTGAAGATTTATGAAATT 1140  
 Qy 1141 GACCAATATGAAGATTTGTTTCTGTCAGGTTTAAAGCTTACCCCAATTCAGATGT 1200  
 Db 1141 GACCAATATGAAGATTTGTTTCTGTCAGGTTTAAAGCTTACCCCAATTCAGATGT 1200  
 Qy 1201 ACGTGGACCTTCTCTCGAAATCAATTTCTTGTGAGCAAAAGGGTCTTGTATACCGGATAC 1260  
 Db 1201 ACGTGGACCTTCTCTCGAAATCAATTTCTTGTGAGCAAAAGGGTCTTGTATACCGGATAC 1260  
 Qy 1261 AGCATATCCAAAGTTTTCGAATCATAAGCACAGCCAGGAGGAATATATATTCATCGAGAA 1320  
 Db 1261 AGCATATCCAAAGTTTTCGAATCATAAGCACAGCCAGGAGGAATATATATTCATCGAGAA 1320

QY 1321 AATGATGATGCCCAATTTTACCAAAATGTTTCAAGTGAATATAGAGGAGAAACCTCAAGTG 1380  
DB |||||  
QY 1321 AATGATGATGCCCAATTTTACCAAAATGTTTCAAGTGAATATAGAGGAGAAACCTCAAGTG 1380  
DB |||||  
QY 1381 CTCGAGAGAGCATCGCAAGTCAAGGCTCTGTTCTCGATGATGATACCATTAACATCT 1440  
DB |||||  
QY 1381 CTCGAGAGAGCATCGCAAGTCAAGGCTCTGTTCTCGATGATGATACCATTAACATCT 1440  
DB |||||  
QY 1441 TGGACCTGGAAAGATGTTTACAGCAAGTCTCCCACTGACAGAGAGATGATACAGAGGA 1500  
DB |||||  
QY 1441 TGGACCTGGAAAGATGTTTACAGCAAGTCTCCCACTGACAGAGAGATGATACAGAGGA 1500  
DB |||||  
QY 1501 GTCTGGAATAGAAAGCTTAACAGAAAGTGTGACAGAGTCTCGAGCAGTACTCTA 1560  
DB |||||  
QY 1501 GTCTGGAATAGAAAGCTTAACAGAAAGTGTGACAGAGTCTCGAGCAGTACTCTA 1560  
DB |||||  
QY 1561 AACATGATGAGCCATTAAGAGGTTCTGTTCAAGTGTGTCATACAAATCCCTTGGC 1620  
DB |||||  
QY 1561 AACATGATGAGCCATTAAGAGGTTCTGTTCAAGTGTGTCATACAAATCCCTTGGC 1620  
DB |||||  
QY 1621 ACATCTGTGAGACATCTTTTAAACTCTCCAGGCTCTCCCTTCAATGCTGTTTAACTCTGTA 1680  
DB |||||  
QY 1621 ACATCTGTGAGACATCTTTTAAACTCTCCAGGCTCTCCCTTCAATGCTGTTTAACTCTGTA 1680  
DB |||||  
QY 1681 ATCTCAATCTATGCAAAATGTTGTTGTTCTCTCTTCAATGCTGTTTAACTCTGTA 1740  
DB |||||  
QY 1681 ATCTCAATCTATGCAAAATGTTGTTGTTCTCTCTTCAATGCTGTTTAACTCTGTA 1740  
DB |||||  
QY 1741 ATTTGTCAAGATGCAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGTAAGTG 1800  
DB |||||  
QY 1741 ATTTGTCAAGATGCAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGTAAGTG 1800  
DB |||||  
QY 1801 ACCGCTCTCAGATATGAGTACTTCTACGTTGATTTTACAGATATGATATGATCTC 1860  
DB |||||  
QY 1801 ACCGCTCTCAGATATGAGTACTTCTACGTTGATTTTACAGATATGATATGATCTC 1860  
DB |||||  
QY 1861 AAATGGGAGTTTCCAGAGAGAAATTTAGAGTTTGGAGAGTACTAGGATCAGGTCCTTTT 1920  
DB |||||  
QY 1861 AAATGGGAGTTTCCAGAGAGAAATTTAGAGTTTGGAGAGTACTAGGATCAGGTCCTTTT 1920  
DB |||||  
QY 1921 GGAAAGTGTATGAACGCAACAGCTTTATGAAATTAGCAAAAAGGAGTCTCAATCCAGTT 1980  
DB |||||  
QY 1921 GGAAAGTGTATGAACGCAACAGCTTTATGAAATTAGCAAAAAGGAGTCTCAATCCAGTT 1980  
DB |||||  
QY 1981 GCGGTCAAAATGCTGAAGAGAAAGAGAGAGTCTGAAAGAGAGGCACTCATGTGAGAA 2040  
DB |||||  
QY 1981 GCGGTCAAAATGCTGAAGAGAAAGAGAGAGTCTGAAAGAGAGGCACTCATGTGAGAA 2040  
DB |||||  
QY 2041 CTCAAGATGATGACCCAGCTGGAGAGCCAGAGATATTTGAACTGTTGATGTTCTCAAC 2100  
DB |||||  
QY 2041 CTCAAGATGATGACCCAGCTGGAGAGCCAGAGATATTTGAACTGTTGATGTTCTCAAC 2100  
DB |||||  
QY 2101 ACATGTCAGGACCAATTTACTGATTTTGAATGTTGATGTTGATGTTGATGTTGATGTTGAT 2160  
DB |||||  
QY 2101 ACATGTCAGGACCAATTTACTGATTTTGAATGTTGATGTTGATGTTGATGTTGATGTTGAT 2160  
DB |||||  
QY 2161 TATCTAAGAGTAAAGAGAGAAATTTTCAAGGCTTGGAGAGAGTCTGAGAGAGAGTCTGAG 2220  
DB |||||  
QY 2161 TATCTAAGAGTAAAGAGAGAAATTTTCAAGGCTTGGAGAGAGTCTGAGAGAGAGTCTGAG 2220  
DB |||||  
QY 2221 AATTTCAAGTTTACCCCACTTTCCCAATCAGATTTTCCAGGCTTCCAGGCTTCCAGGCTTCCAG 2280  
DB |||||  
QY 2221 AATTTCAAGTTTACCCCACTTTCCCAATCAGATTTTCCAGGCTTCCAGGCTTCCAGGCTTCCAG 2280  
DB |||||  
QY 2281 GAAAGTTCAGATACACCGGACTCGGATCAAAATCTCAGGCTTCCAGGCTTCCAGGCTTCCAGG 2340  
DB |||||  
QY 2281 GAAAGTTCAGATACACCGGACTCGGATCAAAATCTCAGGCTTCCAGGCTTCCAGGCTTCCAGG 2340  
DB |||||  
QY 2341 TCTGAAGATGAATTTGAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 2400  
DB |||||  
QY 2341 TCTGAAGATGAATTTGAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 2400  
DB |||||

QY 2401 GTGCTTACATTTGAAGATCTTCTTTGCTTTGCTATATCAAGTTTCCCAAGGATGGAATTT 2460  
DB |||||  
QY 2401 GTGCTTACATTTGAAGATCTTCTTTGCTTTGCTATATCAAGTTTCCCAAGGATGGAATTT 2460  
DB |||||  
QY 2461 CTGGAATTTAAGTCTGTTTCAAGAGACCTTGGCGCCAGGAACTGCTTGTGTCACCCAC 2520  
DB |||||  
QY 2461 CTGGAATTTAAGTCTGTTTCAAGAGACCTTGGCGCCAGGAACTGCTTGTGTCACCCAC 2520  
DB |||||  
QY 2521 GGGAAAGTGGTGAAGATGATGATCTTTGGATTTGGCTCGAGATATCATGATGATGATTCAC 2580  
DB |||||  
QY 2521 GGGAAAGTGGTGAAGATGATGATCTTTGGATTTGGCTCGAGATATCATGATGATGATTCAC 2580  
DB |||||  
QY 2581 TATGTTGTGAGGGCAATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640  
DB |||||  
QY 2581 TATGTTGTGAGGGCAATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640  
DB |||||  
QY 2641 GAAGCATCTACACCATTAAGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700  
DB |||||  
QY 2641 GAAGCATCTACACCATTAAGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700  
DB |||||  
QY 2701 TTCTCACTTGTGTAATCTTACCTGGCATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760  
DB |||||  
QY 2701 TTCTCACTTGTGTAATCTTACCTGGCATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760  
DB |||||  
QY 2761 ATTCAAATGATTTAAATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820  
DB |||||  
QY 2761 ATTCAAATGATTTAAATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820  
DB |||||  
QY 2821 ATGCAATCTCTGCTGGCTTTTGAATCAAGGAAAGGCGCATCTTCCCTTAAATTTGACTTCTG 2880  
DB |||||  
QY 2821 ATGCAATCTCTGCTGGCTTTTGAATCAAGGAAAGGCGCATCTTCCCTTAAATTTGACTTCTG 2880  
DB |||||  
QY 2881 TTTTGAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2940  
DB |||||  
QY 2881 TTTTGAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2940  
DB |||||  
QY 2941 GTTTTCGGAATGCTCTCACCTCAAGGAGGCTTGTAGTTTACCAAAAGGATTTTAAAGT 3000  
DB |||||  
QY 2941 GTTTTCGGAATGCTCTCACCTCAAGGAGGCTTGTAGTTTACCAAAAGGATTTTAAAGT 3000  
DB |||||  
QY 3001 GGGCTACTCTCTCGCGAGGCTCAGGTCGAGATGCTGAGAGGAGGAGGAGGAGGAGGAGGAGG 3060  
DB |||||  
QY 3001 GGGCTACTCTCTCGCGAGGCTCAGGTCGAGATGCTGAGAGGAGGAGGAGGAGGAGGAGGAGG 3060  
DB |||||  
QY 3061 ACTTCACTCCCTCCACCTTATCCCTAAAGGCTGATGATTTACCAAAAGGATTTTAAAGT 3120  
DB |||||  
QY 3061 ACTTCACTCCCTCCACCTTATCCCTAAAGGCTGATGATTTACCAAAAGGATTTTAAAGT 3120  
DB |||||  
QY 3121 CACTTAAAGAGAAATCTTATTAATCAACTGCTGCTTCCAGAGGCTTCTCTAGAGGCTGCT 3180  
DB |||||  
QY 3121 CACTTAAAGAGAAATCTTATTAATCAACTGCTGCTTCCAGAGGCTTCTCTAGAGGCTGCT 3180  
DB |||||  
QY 3181 GCGTTTACTCTTCTTCAAGGAGCTTTTGTAAATCAATCACTCTGCTCAAGGAGGAGGAGGAGG 3240  
DB |||||  
QY 3181 GCGTTTACTCTTCTTCAAGGAGCTTTTGTAAATCAATCACTCTGCTCAAGGAGGAGGAGGAGG 3240  
DB |||||  
QY 3241 GAGGAGCTGATTAATGAACTTTTATTTGAGAGCATGATCTGATCCAGGCTTCTCAGGCGG 3300  
DB |||||  
QY 3241 GAGGAGCTGATTAATGAACTTTTATTTGAGAGCATGATCTGATCCAGGCTTCTCAGGCGG 3300  
DB |||||  
QY 3301 GCTTGAAGTGAATTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3360  
DB |||||  
QY 3301 GCTTGAAGTGAATTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3360  
DB |||||  
QY 3361 ATTTTCTAAGGAGAGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3420  
DB |||||  
QY 3361 ATTTTCTAAGGAGAGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3420  
DB |||||  
QY 3421 TTTTTCAGCTATTTAGTGTATATATTTTATGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3480  
DB |||||  
QY 3421 TTTTTCAGCTATTTAGTGTATATATTTTATGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3480  
DB |||||  
QY 3481 AAAAAAAAAAAAAAAAAAAAAA 3501



QY 1021 AGAAACGACCGGATACACACTGTTTCTCTTCAAAAGCATCCAGTCAATCAGCTTGG 1080  
DB |||||  
QY 1021 AGAAACGACCGGATACACACTGTTTCTCTTCAAAAGCATCCAGTCAATCAGCTTGG 1080  
DB |||||  
QY 1081 GTTACCATCGTAGGAAAGGGAATTTATATAATGCTACCAATTCAGAGTGAAGATTTAGAAAT 1140  
DB |||||  
QY 1081 GTTACCATCGTAGGAAAGGGAATTTATATAATGCTACCAATTCAGAGTGAAGATTTAGAAAT 1140  
DB |||||  
QY 1141 GACCAATATGAAGATTTGTTTTCTGTGCTAGGTTTAAAGCTTACCCACAATCAGATGT 1200  
DB |||||  
QY 1141 GACCAATATGAAGATTTGTTTTCTGTGCTAGGTTTAAAGCTTACCCACAATCAGATGT 1200  
DB |||||  
QY 1201 ACCTGGACCTTCTCTCGAAATATCAATTTCTGTGAGCAAAAGGCTTTGATAACGGATAC 1260  
DB |||||  
QY 1201 ACCTGGACCTTCTCTCGAAATATCAATTTCTGTGAGCAAAAGGCTTTGATAACGGATAC 1260  
DB |||||  
QY 1261 AGCATATCCAAGTTTTCGAATCATPAAGCACACGAGGAGAGATATATATTTCCATGCGAGAA 1320  
DB |||||  
QY 1261 AGCATATCCAAGTTTTCGAATCATPAAGCACACGAGGAGAGATATATATTTCCATGCGAGAA 1320  
DB |||||  
QY 1321 AATGATGATGCCCAATTTACCAAAATGTTTCAAGCTGAATATAAGAGAAACCTCAAGTG 1380  
DB |||||  
QY 1321 AATGATGATGCCCAATTTACCAAAATGTTTCAAGCTGAATATAAGAGAAACCTCAAGTG 1380  
DB |||||  
QY 1381 CTGCGAAGCATCGGCAAGTCTGAGCAAGTCTCCCACTGCGAAGAGAGATACCCATACCATCT 1440  
DB |||||  
QY 1381 CTGCGAAGCATCGGCAAGTCTGAGCAAGTCTCCCACTGCGAAGAGAGATACCCATACCATCT 1440  
DB |||||  
QY 1441 TGGACCTGGAAGAGTGTTCAGACAACTCTCCCACTGCGAAGAGAGATACCAAGAGGA 1500  
DB |||||  
QY 1441 TGGACCTGGAAGAGTGTTCAGACAACTCTCCCACTGCGAAGAGAGATACCAAGAGGA 1500  
DB |||||  
QY 1501 GTCTGGAATAGAAAGGCTTAAAGAGTGTTCAGCAAGTGTTCGAGTGGGTCTGAGCAGTACTCTA 1560  
DB |||||  
QY 1501 GTCTGGAATAGAAAGGCTTAAAGAGTGTTCAGCAAGTGTTCGAGTGGGTCTGAGCAGTACTCTA 1560  
DB |||||  
QY 1561 AACATGAGTGAAGCCATAAAGGTTCTGCTGAGTGTTCGAGTGGGTCTGAGCAGTACTCTA 1620  
DB |||||  
QY 1561 AACATGAGTGAAGCCATAAAGGTTCTGCTGAGTGTTCGAGTGGGTCTGAGCAGTACTCTA 1620  
DB |||||  
QY 1621 ACATCTTGTGAGCAGTCTCTTAAACTCTCCAGGCCCTTCCCTTTCATCCAGAGCAAC 1680  
DB |||||  
QY 1621 ACATCTTGTGAGCAGTCTCTTAAACTCTCCAGGCCCTTCCCTTTCATCCAGAGCAAC 1680  
DB |||||  
QY 1681 ATCTCATTTATGCAACAAATGTTGTTTGTCTCCTCTTCAATGTTGTTTAAACCTGCTA 1740  
DB |||||  
QY 1681 ATCTCATTTATGCAACAAATGTTGTTTGTCTCCTCTTCAATGTTGTTTAAACCTGCTA 1740  
DB |||||  
QY 1741 ATTTGTCAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
DB |||||  
QY 1741 ATTTGTCAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
DB |||||  
QY 1801 ACCGGCTCTCAGATAATGAGTACTTCTACGTTGATTTTCAAGAGATATGATATGATCTC 1860  
DB |||||  
QY 1801 ACCGGCTCTCAGATAATGAGTACTTCTACGTTGATTTTCAAGAGATATGATATGATCTC 1860  
DB |||||  
QY 1861 AAATGGAGTCTTCAAGAGAAATTTAGAGTTTGGAGGTACTAGGATCAGGTGCTTTT 1920  
DB |||||  
QY 1861 AAATGGAGTCTTCAAGAGAAATTTAGAGTTTGGAGGTACTAGGATCAGGTGCTTTT 1920  
DB |||||  
QY 1921 GGAAGTGTATGAAAGCAAGCTTATGAAATTTAGGAAATAGGAGAGTCTCAATCCAGGTT 1980  
DB |||||  
QY 1921 GGAAGTGTATGAAAGCAAGCTTATGAAATTTAGGAAATAGGAGAGTCTCAATCCAGGTT 1980  
DB |||||  
QY 1981 GCGCTCAAAATGCTGAAAGAAAGCAGACAGCTCTGAAAGAGAGCACTCATGTGAGAA 2040  
DB |||||  
QY 1981 GCGCTCAAAATGCTGAAAGAAAGCAGACAGCTCTGAAAGAGAGCACTCATGTGAGAA 2040  
DB |||||  
QY 2041 CTCAGATGATGACCCAGCTGGGAAGCCAGAGATATTTGTGAACCTGCTGGGGCGTGC 2100  
DB |||||  
QY 2041 CTCAGATGATGACCCAGCTGGGAAGCCAGAGATATTTGTGAACCTGCTGGGGCGTGC 2100  
DB |||||

QY 2101 ACACTGTCAGGACCAATTTACTTGAATTTTGAATACTGTTGCTATGCTGATCTTCTCAAC 2160  
DB |||||  
QY 2101 ACACTGTCAGGACCAATTTACTTGAATTTTGAATACTGTTGCTATGCTGATCTTCTCAAC 2160  
DB |||||  
QY 2161 TATCTAAGAAATGAAAGAGAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
DB |||||  
QY 2161 TATCTAAGAAATGAAAGAGAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
DB |||||  
QY 2221 AATTTAGTGTTTTACCCACCTTTTCAATCAATCCAAATTCAGCATGCTGTTCAAGA 2280  
DB |||||  
QY 2221 AATTTAGTGTTTTACCCACCTTTTCAATCAATCCAAATTCAGCATGCTGTTCAAGA 2280  
DB |||||  
QY 2281 GAGTTCAGATACACCCGAGCTCGSATCAAAATCTCAGGCTTCAAGGATTTCAATTTCAAC 2340  
DB |||||  
QY 2281 GAGTTCAGATACACCCGAGCTCGSATCAAAATCTCAGGCTTCAAGGATTTCAATTTCAAC 2340  
DB |||||  
QY 2341 TCTGAAGATGAAATTTGAATATGAAATCCAAAAGGCTGGAAGAGAGAGAGACTTTGAAT 2400  
DB |||||  
QY 2341 TCTGAAGATGAAATTTGAATATGAAATCCAAAAGGCTGGAAGAGAGAGAGACTTTGAAT 2400  
DB |||||  
QY 2401 GTGCTTACATTTGAAGATCTTCTTGTCTTGCATATCAAGTTGCCAAAGGAATGGAATTT 2460  
DB |||||  
QY 2401 GTGCTTACATTTGAAGATCTTCTTGTCTTGCATATCAAGTTGCCAAAGGAATGGAATTT 2460  
DB |||||  
QY 2461 CTGGAATTTAAGTCTGCTGTTTCAAGAGACTGCGCCAGAGAGCTGTTGTCAACCCAC 2520  
DB |||||  
QY 2461 CTGGAATTTAAGTCTGCTGTTTCAAGAGACTGCGCCAGAGAGCTGTTGTCAACCCAC 2520  
DB |||||  
QY 2521 GGGAAAGTGTGGAAGATATGAGTCTTGGATGCTCGAGATATCATGAGTGTATTTCCAAAC 2580  
DB |||||  
QY 2521 GGGAAAGTGTGGAAGATATGAGTCTTGGATGCTCGAGATATCATGAGTGTATTTCCAAAC 2580  
DB |||||  
QY 2581 TATGTTTCTCAGGGGCAATGCCCCTGCTGCTGTAATAATGGAATGGCCCCGAAAGCCCTGTTT 2640  
DB |||||  
QY 2581 TATGTTTCTCAGGGGCAATGCCCCTGCTGCTGTAATAATGGAATGGCCCCGAAAGCCCTGTTT 2640  
DB |||||  
QY 2641 GAAGGCAATCTACACATTAAGAGTGTCTGCTGATATGGAATATTTACTGTGGAAATC 2700  
DB |||||  
QY 2641 GAAGGCAATCTACACATTAAGAGTGTCTGCTGATATGGAATATTTACTGTGGAAATC 2700  
DB |||||  
QY 2701 TTCTCAGTCTGCTGGAATCTTACCTCGGATTCGGTGTGATGCTTCTACAAATCTG 2760  
DB |||||  
QY 2701 TTCTCAGTCTGCTGGAATCTTACCTCGGATTCGGTGTGATGCTTCTACAAATCTG 2760  
DB |||||  
QY 2761 ATTCAAAATGGAATTTAAATGGAATCAGCCATTTTATGCTACAGAGAAATATACATTATA 2820  
DB |||||  
QY 2761 ATTCAAAATGGAATTTAAATGGAATCAGCCATTTTATGCTACAGAGAAATATACATTATA 2820  
DB |||||  
QY 2821 ATGCAATCTGCTGGGCTTTTGAATCAAGGAAACGGCCATCTTCCCTAAATTTGACTTCG 2880  
DB |||||  
QY 2821 ATGCAATCTGCTGGGCTTTTGAATCAAGGAAACGGCCATCTTCCCTAAATTTGACTTCG 2880  
DB |||||  
QY 2881 TTTTGTAGGATGCTCAGTGGCAGATGAGAGAGAGATGATACAGATGTTGGATGGCGGT 2940  
DB |||||  
QY 2881 TTTTGTAGGATGCTCAGTGGCAGATGAGAGAGAGATGATACAGATGTTGGATGGCGGT 2940  
DB |||||  
QY 2941 GTTTCGGAATGCTCTCAGCCTTACCAAAACAGCGACTTTTACGACAGAGATGGAATTTG 3000  
DB |||||  
QY 2941 GTTTCGGAATGCTCTCAGCCTTACCAAAACAGCGACTTTTACGACAGAGATGGAATTTG 3000  
DB |||||  
QY 3001 GGGCTACTCTCTCCGAGGCTCAGGTGCGAATTTGCTAGAGGAAATTTAGTTTTAAGG 3060  
DB |||||  
QY 3061 ACTTCATCTCTCCACCTATCCCTAAAGGCTGTAGATTTACCAAAACAGATTTAATTTCAAT 3120  
DB |||||  
QY 3121 CACTAAAAGAAATCTATTATCACTGCTTCAACGACTTTTCTCTAGAGAGCCGCTCT 3180  
DB |||||  
QY 3121 CACTAAAAGAAATCTATTATCACTGCTTCAACGACTTTTCTCTAGAGAGCCGCTCT 3180  
DB |||||  
QY 3181 GCGTTTACTCTTTTTCAAAGGGAATTTTGTAAAATCAAAATCATCTCTGTCACAAAGGCGAG 3240  
DB |||||

Db 3181 GCCTTTACTCTGTTTCAAGGAGCTTTTGTAAATCAATCATCTGTACAAAGCAG 3240  
 QY 3241 GAGGAGCTGATAATGAACCTTTTATTTGGAGCATTGATCTGCATCCAAAGCCTTCTCAGGCCG 3300  
 Db 3241 GAGGAGCTGATAATGAACCTTTTATTTGGAGCATTGATCTGCATCCAAAGCCTTCTCAGGCCG 3300  
 QY 3301 GCCTTGAGTGAATTTGTTACCTGAAGTACAGTATATTTCTTGAATACATAAAACAAAGC 3360  
 Db 3301 GCCTTGAGTGAATTTGTTACCTGAAGTACAGTATATTTCTTGAATACATAAAACAAAGC 3360  
 QY 3361 ATTTTGTAGGAGAGCTAATATGATTTTAAAGTCTATGTTTAAATTAATATGTAATA 3420  
 Db 3361 ATTTTGTAGGAGAGCTAATATGATTTTAAAGTCTATGTTTAAATTAATATGTAATA 3420  
 QY 3421 TTTTTCAGCTATTAGTGATATATTTTATGGTGCGGAATATAATTTCTACTACAGAAAAA 3480  
 Db 3421 TTTTTCAGCTATTAGTGATATATTTTATGGTGCGGAATATAATTTCTACTACAGAAAAA 3480  
 QY 3481 AAAAAAAAAAAAAAAAAAAAAA 3501  
 Db 3481 AAAAAAAAAAAAAAAAAAAAAA 3501

## RESULT 5

AAAX77515  
 ID AAX77515 standard; cDNA; 3501 BP.

XX AC AAX77515;

XX DT 05-AUG-1999 (first entry)

XX DE Human flk-2 cDNA.

XX KW Human; flk-2; flk-1; cell isolation; fetal liver kinase; receptor;  
 XX monoclinal; polyclonal; antibody; tyrosine kinase; db.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX CDS 58..3039

XX FT /\*tag= a

XX FT /product= "flk-2"

XX US5912133-A.

XX PD 15-JUN-1999.

XX PF 10-FEB-1998; 98US-0021324.

XX PR 19-NOV-1992; 92US-0977451.

XX PR 02-APR-1991; 91US-0679666.

XX PR 28-JUN-1991; 91US-0728913.

XX PR 15-NOV-1991; 91US-0793065.

XX PR 24-DEC-1991; 91US-0813593.

XX PR 26-JUN-1992; 92US-0906397.

XX PR 12-NOV-1992; 92US-0975049.

XX PR 30-APR-1993; 93US-0055269.

XX PR 31-OCT-1994; 94US-0252498.

XX PR 15-FEB-1996; 96US-0601891.

XX (UYPB-) UNIV PRINCETON.

XX PI Lemischka IR;

XX DR WPI; 1999-357194/30.

XX DR P-PSDB; AAY08617.

XX PT Isolating hematopoietic cells expressing fetal liver kinase 1  
 XX receptors

XX PS Disclosure; Fig 1b; 59pp; English.

XX XX

CC This invention describes a novel method of isolating cells expressing  
 CC fetal liver kinase 1 (flk-1) receptors on their surface and comprises  
 CC binding the cells to a polyclonal or monoclonal antibody specific to  
 CC the flk-1 receptor and isolating the cells that have bound to the  
 CC antibody. The method can be used to isolate hematopoietic stem cells in  
 CC any mammal but preferably a rat, mouse, rabbit or human. The proteins of  
 CC the invention belong to the receptor protein family. This sequence  
 CC encodes the human flk-2 protein which is used in the method of the  
 CC invention.

SQ Sequence 3501 BP; 1068 A; 709 C; 784 G; 940 T; 0 other;

Query Match 100.0%; Score 3501; DB 20; Length 3501;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGAGGGCGCATCCGAGGGCTGGCGCGCGCCCTGGGGGACCCCGGGCTCCGGAGGCCATG 60

Db 1 CGAGGGCGCATCCGAGGGCTGGCGCGCGCCCTGGGGGACCCCGGGCTCCGGAGGCCATG 60

QY 61 CCGGCGTTGGCGCGCGACGCGCGCACCGTCCGCTGCTGTTGTTTTCTGCAATGATA 120

Db 61 CCGGCGTTGGCGCGCGACGCGCGCACCGTCCGCTGCTGTTGTTTTCTGCAATGATA 120

QY 121 TTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTTAATCAATAAG 180

Db 121 TTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTTAATCAATAAG 180

QY 181 AACAAATGATTCATCAGTGGGGAAGTCAATCATATATCCCATGGTATCAGAAATCCCGGAA 240

Db 181 AACAAATGATTCATCAGTGGGGAAGTCAATCATATATCCCATGGTATCAGAAATCCCGGAA 240

QY 241 GACCTGGGTGCTGTTTGAGACCCCGAGAGCTCAGGGACAGTGTACGAAGCTGCCGCTGTG 300

Db 241 GACCTGGGTGCTGTTTGAGACCCCGAGAGCTCAGGGACAGTGTACGAAGCTGCCGCTGTG 300

QY 301 GAAGTGGAGTATCTGCTTCCATCACACTGCAAGTGTGTCGATGCCCGAGGGAACATT 360

Db 301 GAAGTGGAGTATCTGCTTCCATCACACTGCAAGTGTGTCGATGCCCGAGGGAACATT 360

QY 361 TCCTGTCTGGGTCTTTAAGCACAGCTCCCTGAATGCCAGCCACATTTTGTATTACAA 420

Db 361 TCCTGTCTGGGTCTTTAAGCACAGCTCCCTGAATGCCAGCCACATTTTGTATTACAA 420

QY 421 AACAGAGGAGTGTTCCTCATGTTTCATTTTCAAAATGACAGAAACCCCAAGCTGGAGATAC 480

Db 421 AACAGAGGAGTGTTCCTCATGTTTCATTTTCAAAATGACAGAAACCCCAAGCTGGAGATAC 480

QY 481 CTACTTTTATTTCAGAGTGAAGCTACCAATTACACAATATTGTTTACAGTGAGTATAAGA 540

Db 481 CTACTTTTATTTCAGAGTGAAGCTACCAATTACACAATATTGTTTACAGTGAGTATAAGA 540

QY 541 AATACCTCTGTTTACACATTAAAGAACCTTACTTTTAGAAAAATGGAAAAACGAGAGGCC 600

Db 541 AATACCTCTGTTTACACATTAAAGAACCTTACTTTTAGAAAAATGGAAAAACGAGAGGCC 600

QY 601 CTGCTCTGCATATCTGAGCGGTTCCAGAGCCGATCGTGAATGGGTGCTTTCGATTCGA 660

Db 601 CTGCTCTGCATATCTGAGCGGTTCCAGAGCCGATCGTGAATGGGTGCTTTCGATTCGA 660

QY 661 CAGGGGGAAGCTGTAAAGAAAGAACGTCAGCTGTTTGTAAAGAGGAGAAAAAGTGCCT 720

Db 661 CAGGGGGAAGCTGTAAAGAAAGAACGTCAGCTGTTTGTAAAGAGGAGAAAAAGTGCCT 720

QY 721 CATGAATTTATTTGGGACGGAATAAGGTGCTGTGCCAGAAATGAATCTGGGAGGAAATGC 780

Db 721 CATGAATTTATTTGGGACGGAATAAGGTGCTGTGCCAGAAATGAATCTGGGAGGAAATGC 780

QY 781 ACCAGGCTGTTTCAATAGATCTAAATCAAACTCTCAGACCAATTGCCCAATTTT 840

Db 781 ACCAGGCTGTTTCAATAGATCTAAATCAAACTCTCAGACCAATTGCCCAATTTT 840

QY 841 CTTAAAGTAGGGGAAACCTTTATGATAAGGTGCAAGCTGTTTCATGTGTGAACCATGGATT 900

Db 841 CTTAAAGTAGGGAAACCTTTATGATGAAGTGGCAAGCTGTTTCATGTGAACCATGATTC 900  
QY 901 GGCGTCACTGGGAAATTAGAAAAACAAGCACTCGAGGAGGCAACTACTTTGAGATGAGT 960  
Db 901 GGCGTCACTGGGAAATTAGAAAAACAAGCACTCGAGGAGGCAACTACTTTGAGATGAGT 960  
QY 961 ACCTATTCAACAACAGAACTATGATACGGATCTCTGTTTCTGTTGATCATCATGAGGCA 1020  
Db 961 ACCTATTCAACAACAGAACTATGATACGGATCTCTGTTTCTGTTGATCATCATGAGGCA 1020  
QY 1021 AGAAACGACACCGGATACACCTTTGTTCTCTTCAAGCATCCAGCTCAATCAGCTTTG 1080  
Db 1021 AGAAACGACACCGGATACACCTTTGTTCTCTTCAAGCATCCAGCTCAATCAGCTTTG 1080  
QY 1081 GTTACCATCTAGAAAGGATTTAATAATGCTACCAATTCAGAGTGAAGATTTAGAAAT 1140  
Db 1081 GTTACCATCTAGAAAGGATTTAATAATGCTACCAATTCAGAGTGAAGATTTAGAAAT 1140  
QY 1141 GACCAATATGAAGATTTGTTTCTGTCAGGTTTAAAGCCTACCCACAATCAGATGT 1200  
Db 1141 GACCAATATGAAGATTTGTTTCTGTCAGGTTTAAAGCCTACCCACAATCAGATGT 1200  
QY 1201 ACGTGGACCTTCTCTCGAAATCAATTTCTGTCAGGTTTAAAGCCTACCCACAATCAGATGT 1260  
Db 1201 ACGTGGACCTTCTCTCGAAATCAATTTCTGTCAGGTTTAAAGCCTACCCACAATCAGATGT 1260  
QY 1261 AGCATATCCAAGTTTGGCAATCAATACACACCGCAGGAGATATATATTTCCATGCGAA 1320  
Db 1261 AGCATATCCAAGTTTGGCAATCAATACACACCGCAGGAGATATATATTTCCATGCGAA 1320  
QY 1321 AATGATGATGCCCAATTTACCAAAATGTTTCAAGCTGAAATATATGAGAGAAACCTCAAGT 1380  
Db 1321 AATGATGATGCCCAATTTACCAAAATGTTTCAAGCTGAAATATATGAGAGAAACCTCAAGT 1380  
QY 1381 CTGCGAAGAGCATCGCAAGTCAAGGCTGCTGTTTCTCGAGTGAATACCAATACCATCT 1440  
Db 1381 CTGCGAAGAGCATCGCAAGTCAAGGCTGCTGTTTCTCGAGTGAATACCAATACCATCT 1440  
QY 1441 TGGACCTGGAAGATGTTTCAAGATCTCCCAACTGCAAGAGAGATCAAGAGGA 1500  
Db 1441 TGGACCTGGAAGATGTTTCAAGATCTCCCAACTGCAAGAGAGATCAAGAGGA 1500  
QY 1501 GTCTGAAATAGAAAGCTTAACAGAAAGTGTGTTGGAAGTGTGAGAGTCACTCTA 1560  
Db 1501 GTCTGAAATAGAAAGCTTAACAGAAAGTGTGTTGGAAGTGTGAGAGTCACTCTA 1560  
QY 1561 AACATGAGTGAAGCCATAAAGGTTTCTGTTCAAGTGTGTCATACATTTCCCTTGGC 1620  
Db 1561 AACATGAGTGAAGCCATAAAGGTTTCTGTTCAAGTGTGTCATACATTTCCCTTGGC 1620  
QY 1621 ACATCTTGTGAGACGATCCCTTTTAAACTCTCCAGGCCCCCTTCCCTTTCATCCAAAGCAAC 1680  
Db 1621 ACATCTTGTGAGACGATCCCTTTTAAACTCTCCAGGCCCCCTTCCCTTTCATCCAAAGCAAC 1680  
QY 1681 ATCTCATTTCTAGCAACAAATGTTGTTGTTCTCTCTTCAATGCTGTTTAAACCTGCTA 1740  
Db 1681 ATCTCATTTCTAGCAACAAATGTTGTTGTTGTTCTCTCTCTTCAATGCTGTTTAAACCTGCTA 1740  
QY 1741 ATTTGTCAAGTACAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGTTGAGGTG 1800  
Db 1741 ATTTGTCAAGTACAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGTTGAGGTG 1800  
QY 1801 ACCGGCTCTCAGATAATGAGTACTTCTAGTTGATTTTCAAGAAATATGATATCTC 1860  
Db 1801 ACCGGCTCTCAGATAATGAGTACTTCTAGTTGATTTTCAAGAAATATGATATCTC 1860  
QY 1861 AAATGGAGTTTCCAAAGAAAAATTTAGATTTGGAAAGTACTAGGATCAGGTGCTTTT 1920  
Db 1861 AAATGGAGTTTCCAAAGAAAAATTTAGATTTGGAAAGTACTAGGATCAGGTGCTTTT 1920  
QY 1921 GGAAGAGTGAACCGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGTT 1980

Db 1921 GGAAGAGTGAACCGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGTT 1980  
QY 1981 GGCGTCAAAATGCTGAAAGAAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCAGAA 2040  
Db 1981 GGCGTCAAAATGCTGAAAGAAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCAGAA 2040  
QY 2041 CTCAAGATGATGACCCAGCTGGGAAGCCACGAGAAATTTGGAACCTGCTGGGGCGTGC 2100  
Db 2041 CTCAAGATGATGACCCAGCTGGGAAGCCACGAGAAATTTGGAACCTGCTGGGGCGTGC 2100  
QY 2101 ACATGTCAGGACCAATTTACTTGTATTTTGAATFACCTGTTGCTATGTTGATCTTCTCAAC 2160  
Db 2101 ACATGTCAGGACCAATTTACTTGTATTTTGAATFACCTGTTGCTATGTTGATCTTCTCAAC 2160  
QY 2161 TATCTAAGAGTAAAGAGAAAAATTTCAAGGACTTTGGAAGAGATTTTCAAGGAACAC 2220  
Db 2161 TATCTAAGAGTAAAGAGAAAAATTTCAAGGACTTTGGAAGAGATTTTCAAGGAACAC 2220  
QY 2221 AATTTCAAGTTTACCCCACTTTTCCAATCAGATCCAAATTTCCAGCATGCTGTTCAAGA 2280  
Db 2221 AATTTCAAGTTTACCCCACTTTTCCAATCAGATCCAAATTTCCAGCATGCTGTTCAAGA 2280  
QY 2281 GAAGTTTCAGATACACCCGGAACCTCGGATCAAAATCTCAGGGCTTTCATGGAAATTCATTTCC 2340  
Db 2281 GAAGTTTCAGATACACCCGGAACCTCGGATCAAAATCTCAGGGCTTTCATGGAAATTCATTTCC 2340  
QY 2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAAAGGCTGGAAAGAGAGGAGGACTTGAAT 2400  
Db 2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAAAGGCTGGAAAGAGAGGAGGACTTGAAT 2400  
QY 2401 GTGCTTACATTTGAAGATCTTCTTGTGTCATATCAAGTTGCTCAAGAGAAATGGAATTT 2460  
Db 2401 GTGCTTACATTTGAAGATCTTCTTGTGTCATATCAAGTTGCTCAAGAGAAATGGAATTT 2460  
QY 2461 CTGGAATTTAAGTCGTGTTTCAAGAGACCTGGCCGCAAGAGACGTGTTGCAACCCAC 2520  
Db 2461 CTGGAATTTAAGTCGTGTTTCAAGAGACCTGGCCGCAAGAGACGTGTTGCAACCCAC 2520  
QY 2521 GGAAAGTGTGGAAGATATGATGATTTGGATTTGGCTGAGATATCATGAGTATTTCCAAC 2580  
Db 2521 GGAAAGTGTGGAAGATATGATGATTTGGATTTGGCTGAGATATCATGAGTATTTCCAAC 2580  
QY 2581 TATGTTGTGAGGGCAATGCCGTCTGCCCTGTAAAAATGGAATGGCCCCGAAAGCTGTTT 2640  
Db 2581 TATGTTGTGAGGGCAATGCCGTCTGCCCTGTAAAAATGGAATGGCCCCGAAAGCTGTTT 2640  
QY 2641 GAAGCATCTACACCAATTAAGAGTGTGTCATATGGAATATTTACTGTGGGAAATC 2700  
Db 2641 GAAGCATCTACACCAATTAAGAGTGTGTCATATGGAATATTTACTGTGGGAAATC 2700  
QY 2701 TTCTCACTTGGTGAATTCCTTACCTGGCATTCGGTTCAGTAACTTCTACAAATG 2760  
Db 2701 TTCTCACTTGGTGAATTCCTTACCTGGCATTCGGTTCAGTAACTTCTACAAATG 2760  
QY 2761 ATTCAAAATGGAATTTAAATGGAATCAGCCATTTTATGCTACAGAAATAATACATTATA 2820  
Db 2761 ATTCAAAATGGAATTTAAATGGAATCAGCCATTTTATGCTACAGAAATAATACATTATA 2820  
QY 2821 ATGCAATCTCTGCTGGGCTTTTGAATCAAGAAACGGCCATCTTCCCTAATTTGACTTCG 2880  
Db 2821 ATGCAATCTCTGCTGGGCTTTTGAATCAAGAAACGGCCATCTTCCCTAATTTGACTTCG 2880  
QY 2881 TTTTATGAGATGTCAGCTGGCAGATGAGAGAGCGATGATATCAGATGTTGGATGCGCT 2940  
Db 2881 TTTTATGAGATGTCAGCTGGCAGATGAGAGAGCGATGATATCAGATGTTGGATGCGCT 2940  
QY 2941 GTTTCGGAATGTCCTCACACCTTACCAAAACAGGCGACCTTTTTCAGCAGAGAGATGGAATTTG 3000  
Db 2941 GTTTCGGAATGTCCTCACACCTTACCAAAACAGGCGACCTTTTTCAGCAGAGAGATGGAATTTG 3000  
QY 3001 GGGCTACTCTCTCCGAGGCTCAGGTGGAAGATTCGTAGAGGAACAATTTAGTTTAAAG 3060  
Db 3001 GGGCTACTCTCTCCGAGGCTCAGGTGGAAGATTCGTAGAGGAACAATTTAGTTTAAAG 3060



QY 3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACAAACCAAGATTAAATTTTCAT 3120  
 DB |||||  
 QY 3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACAAACCAAGATTAAATTTTCAT 3120  
 DB |||||  
 QY 3121 CACTAAAGAAATCTATTATCACTGCTCTTCCACAGACTTTTCTCTAGAGCCGCTCT 3180  
 DB |||||  
 QY 3121 CACTAAAGAAATCTATTATCACTGCTCTTCCACAGACTTTTCTCTAGAGCCGCTCT 3180  
 DB |||||  
 QY 3181 GCGTTTACTCTTGTCTTCAAGGAGCTTTTGTAAATCAATCATCTGTCAACAGGCAG 3240  
 DB |||||  
 QY 3181 GCGTTTACTCTTGTCTTCAAGGAGCTTTTGTAAATCAATCATCTGTCAACAGGCAG 3240  
 DB |||||  
 QY 3241 GAGGAGCTGATAATGAATTTATTTGGAGCATTTGATCTGCATCCAAAGCCCTTCTCAGGCCG 3300  
 DB |||||  
 QY 3241 GAGGAGCTGATAATGAATTTATTTGGAGCATTTGATCTGCATCCAAAGCCCTTCTCAGGCCG 3300  
 DB |||||  
 QY 3301 GCTTGAGTGAATTTGACCTGAGTACAGTATATCTTCTTAATACATAAACAAGC 3360  
 DB |||||  
 QY 3301 GCTTGAGTGAATTTGACCTGAGTACAGTATATCTTCTTAATACATAAACAAGC 3360  
 DB |||||  
 QY 3361 ATTCTGCTAAGGAGAGCTAATATGATTTTAAAGTCTATGTTTAAATTAATATGATAA 3420  
 DB |||||  
 QY 3361 ATTCTGCTAAGGAGAGCTAATATGATTTTAAAGTCTATGTTTAAATTAATATGATAA 3420  
 DB |||||  
 QY 3421 TTTTTCAGCTATTAGTGATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAAA 3480  
 DB |||||  
 QY 3421 TTTTTCAGCTATTAGTGATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAAA 3480  
 DB |||||  
 QY 3481 AAAAAAAAAAAAAAAAAAAAAA 3501  
 DB AAAAAAAAAAAAAAAAAAAAAA 3501

## RESULT 6

ABS55045  
 ID ABS55045 standard; cDNA; 3501 BP.  
 AC ABS55045;  
 XX  
 DT 09-DEC-2002 (first entry)  
 XX  
 DE Human cDNA encoding receptor protein tyrosine kinase, FLK-2.  
 KW Human; ss; gene; FLK-2; foetal liver kinase; antianaemic;  
 KW ophthalmological; receptor protein tyrosine kinase; aplastic anaemia;  
 KW primitive haematopoietic cell; stem cell; macrocytic anaemia;  
 KW bone marrow damage; cancer chemotherapy.  
 XX  
 OS Homo sapiens.  
 PH  
 FT Key Location/Qualifiers  
 FT misc\_feature 1..1689  
 FT /tag= a  
 FT /note= "This region is specifically claimed in claim 72"  
 CDS 58..3039  
 FT /tag= b  
 FT /product= "FLK-2"  
 FT /note= "This CDS (minus the stop codon) is  
 specifically claimed in claim 73"  
 sig\_peptide 58..138  
 FT /tag= c  
 FT mat\_peptide 139..3036  
 FT /tag= d  
 FT /label= "Mature FLK 2"  
 FT /note= "This region is specifically claimed in claim 74"  
 XX  
 PN US2002119545-A1.  
 XX  
 PD 29-AUG-2002.  
 XX  
 PF 01-JUN-2001; 2001US-0872136.  
 XX

PR 19-NOV-1992; 92US-0977451.  
 PR 30-APR-1993; 93US-0055269.  
 PR 31-OCT-1994; 94US-0252498.  
 PR 15-FEB-1996; 96US-0601891.  
 PR 10-FEB-1998; 98US-0021324.  
 PR 10-DEC-1998; 98US-0208786.  
 PR 02-APR-1991; 91US-0679666.  
 PR 28-JUN-1991; 91US-0728913.  
 PR 15-NOV-1991; 91US-0793065.  
 PR 24-DEC-1991; 91US-0813593.  
 XX (LEMI/) LEMISCHKA I R.  
 XX  
 XX Lemischka IR;  
 XX WPI; 2002-731356/79.  
 DR P-PSDB; ABG70916.  
 XX  
 XX New protein tyrosine kinase expressed in primitive hematopoietic cells  
 PT (HC) and not expressed in mature HC, and ligands for the protein,  
 PT useful for stimulating proliferation of primitive hematopoietic stem  
 PT cells  
 XX  
 PS Claim 9; Fig 1b; 64pp; English.  
 XX  
 CC The invention relates to a receptor protein tyrosine kinase (pTK)  
 CC expressed in primitive haematopoietic cells and not expressed in mature  
 CC haematopoietic cells, and named FLK-1 and -2 (foetal liver kinase).  
 CC Also included are the nucleic acids encoding the FLK proteins, FLK  
 CC expression vectors, a ligand that binds to human or murine FLK-2 or FLK-1  
 CC and stimulates the proliferation and/or differentiation of the primitive  
 CC haematopoietic cells and a murine cell line 2018 having American Type  
 CC Culture Collection (ATCC) accession number ATCC CRL 10907. The ligands  
 CC are useful for stimulating the proliferation and/or differentiation of  
 CC primitive mammalian haematopoietic stem cells. The receptor pTK molecules  
 CC are useful for stimulating the self renewal of the totipotent  
 CC haematopoietic stem cell and to stimulate the development of all cells of  
 CC the haematopoietic system both in vitro and in vivo. The ligands for the  
 CC receptors act as haematopoietic growth factors. The ligands are useful in  
 CC treating humans whose primitive stem cells do not sufficiently undergo  
 CC self-renewal e.g. macrocytic and aplastic anaemia and bone marrow damage  
 CC resulting from cancer chemotherapy and radiation. The present  
 CC sequence encodes human FLK-2.  
 XX  
 SQ Sequence 3501 BP; 1068 A; 709 C; 784 G; 940 T; 0 other;  
 Query Match 100.0%; Score 3501; DB 24; Length 3501;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGAGGCGGCATCCGAGGCTGGGCGGCGGCCCTGGGGGACCCCGGGCTCCGAGGCCATG 60  
 DB |||||  
 QY 1 CGAGGCGGCATCCGAGGCTGGGCGGCGGCCCTGGGGGACCCCGGGCTCCGAGGCCATG 60  
 DB |||||  
 QY 61 CGGCGTGTGGCGCGGACCGGCGGCGCACCGTGCCTGCTGTTTCTGCAATGATA 120  
 DB |||||  
 QY 61 CGGCGTGTGGCGCGGACCGGCGGCGCACCGTGCCTGCTGTTTCTGCAATGATA 120  
 DB |||||  
 QY 121 TTTTGGGACTATTACAATCAAGTCTGCTGCTGATCAAGTGTGTTTAAATCAATCAATG 180  
 DB |||||  
 QY 121 TTTTGGGACTATTACAATCAAGTCTGCTGCTGATCAAGTGTGTTTAAATCAATCAATG 180  
 DB |||||  
 QY 181 AACAAATGATTCATCAGTGGGAGTCAATCATCATATCCATGCTATCAGAAATCCCGGAA 240  
 DB |||||  
 QY 181 AACAAATGATTCATCAGTGGGAGTCAATCATCATATCCATGCTATCAGAAATCCCGGAA 240  
 DB |||||  
 QY 241 GACCTCGGCTGTGCGTTGAGACCCAGAGCTCAGGACAGTGTACGAAGCTGCCGCTGTG 300  
 DB |||||  
 QY 241 GACCTCGGCTGTGCGTTGAGACCCAGAGCTCAGGACAGTGTACGAAGCTGCCGCTGTG 300  
 DB |||||  
 QY 301 GAAGTGAATGATCTGCTTCCATCACAATGCTGCTGATGCCCGGAGCAATT 360  
 DB |||||  
 QY 301 GAAGTGAATGATCTGCTTCCATCACAATGCTGCTGATGCCCGGAGCAATT 360  
 DB |||||



QY 361 TCCTGTCTCTGGTCTTTAAGCACAGCTCCCTGAAATTCGACGACCATTTTGATTTACAA 420  
DB 361 TCCTGTCTCTGGTCTTTAAGCACAGCTCCCTGAAATTCGACGACCATTTTGATTTACAA 420  
QY 421 AACACAGGAGTGTCTTCCATGTGTCAATTTGAAAAATGACAGAAACCCCAAGCTGAGAAATAC 480  
DB 421 AACACAGGAGTGTCTTCCATGTGTCAATTTGAAAAATGACAGAAACCCCAAGCTGAGAAATAC 480  
QY 481 CTACTTTTTATTACAGAGTGAAGCTACAAATACAAATATGTTTACAGTGAATTAAGA 540  
DB 481 CTACTTTTTATTACAGAGTGAAGCTACAAATACAAATATGTTTACAGTGAATTAAGA 540  
QY 541 AATACCTGCTTTACACATTTAGAACACCTTACTTTAGAAAAATGAAAAACCCAGGACGCC 600  
DB 541 AATACCTGCTTTACACATTTAGAACACCTTACTTTAGAAAAATGAAAAACCCAGGACGCC 600  
QY 601 CTGGTCTGCATATCTGAGACGCTTCAGAGCCGATCGTGGAAATGGGTGCTTTCGGATTC 660  
DB 601 CTGGTCTGCATATCTGAGAGGCTTCAGAGCCGATCGTGGAAATGGGTGCTTTCGGATTC 660  
QY 661 CAGGGGAAAGCTGTAAAGAGAAAGTCCAGCTGTGTTTAAAGAGGAGGAAAGTGCCT 720  
DB 661 CAGGGGAAAGCTGTAAAGAGAAAGTCCAGCTGTGTTTAAAGAGGAGGAAAGTGCCT 720  
QY 721 CATGAATTTATTTGGGACGACATAAGGTGCTGTGCCAGAAATGAACTGGGCGAGGAATGC 780  
DB 721 CATGAATTTATTTGGGACGACATAAGGTGCTGTGCCAGAAATGAACTGGGCGAGGAATGC 780  
QY 781 ACCAGGCTGTTCACAAATAGATCTAAATCAAACTCCTCAGACCAATTCGCCACAATTTATTT 840  
DB 781 ACCAGGCTGTTCACAAATAGATCTAAATCAAACTCCTCAGACCAATTCGCCACAATTTATTT 840  
QY 841 CTTAAAGTAGGGGAAACCTTTATGGATAAGGTGCAAAAGCTGTTCAATGTGAACCATGATTC 900  
DB 841 CTTAAAGTAGGGGAAACCTTTATGGATAAGGTGCAAAAGCTGTTCAATGTGAACCATGATTC 900  
QY 901 GGGCTCACTGGGAATTTAGAAAACAAAGCACTCGAGGAGGCAACTTACTTTGAGATGAGT 960  
DB 901 GGGCTCACTGGGAATTTAGAAAACAAAGCACTCGAGGAGGCAACTTACTTTGAGATGAGT 960  
QY 961 ACCTATTCACAAACAGAACTATGATACGGATTCCTGTTGCTTTGATCATCAGTGGCA 1020  
DB 961 ACCTATTCACAAACAGAACTATGATACGGATTCCTGTTGCTTTGATCATCAGTGGCA 1020  
QY 1021 AGAAACGACACCGGATCTACTGTTCTCTTTCAAAGCATCCCAAGTCAATCAGCTTTG 1080  
DB 1021 AGAAACGACACCGGATCTACTGTTCTCTTTCAAAGCATCCCAAGTCAATCAGCTTTG 1080  
QY 1081 GTTACCATCGTAGGAAGGGAATTAATAATGCTTACCAATTCAGTGAAGATTAATAAT 1140  
DB 1081 GTTACCATCGTAGGAAGGGAATTAATAATGCTTACCAATTCAGTGAAGATTAATAAT 1140  
QY 1141 GACCAATATGAAGATTTTGTCTGTCTGTCAGTGTAAAGCTTACCCACCAATCAGATGT 1200  
DB 1141 GACCAATATGAAGATTTTGTCTGTCTGTCAGTGTAAAGCTTACCCACCAATCAGATGT 1200  
QY 1201 AGGTGACCTCTCTCGAAAATCATTTCTTTGTGAGCAAAAGGCTCTGATACGGATAC 1260  
DB 1201 AGGTGACCTCTCTCGAAAATCATTTCTTTGTGAGCAAAAGGCTCTGATACGGATAC 1260  
QY 1261 AGCATATCCAAGTTTTCGAATCATAGACCAAGCCAGGAGGAATATATATTTCCATGCGAAA 1320  
DB 1261 AGCATATCCAAGTTTTCGAATCATAGACCAAGCCAGGAGGAATATATATTTCCATGCGAAA 1320  
QY 1321 AATGATATGCCAATTTACCAAAATGTTCAAGCTGAATATAAGAGGAAACCTCAAGTG 1380  
DB 1321 AATGATATGCCAATTTACCAAAATGTTCAAGCTGAATATAAGAGGAAACCTCAAGTG 1380  
QY 1381 CTCGAGAAAGCATCGCAAGTCAGGCTCTGTTTCTCGGATGGATACCCATTTACCATCT 1440  
DB 1381 CTCGAGAAAGCATCGCAAGTCAGGCTCTGTTTCTCGGATGGATACCCATTTACCATCT 1440

QY 1441 TGGACCTGGAAGAGTGTTCAGACAAGTCTCCCAACTGCACAGAGATCACAGAAGGA 1500  
DB 1441 TGGACCTGGAAGAGTGTTCAGACAAGTCTCCCAACTGCACAGAGATCACAGAAGGA 1500  
QY 1501 GTCTGGAATAGAAAGGCTAACAGAAAGTGTTCGACAGTGGGTGTCGAGCAGTACTCTA 1560  
DB 1501 GTCTGGAATAGAAAGGCTAACAGAAAGTGTTCGACAGTGGGTGTCGAGCAGTACTCTA 1560  
QY 1561 AACATGAGTGAAGCCATAAAAGGTTCTGTCAGTCAAGTGTGTCATACAAATTCCTTTGGC 1620  
DB 1561 AACATGAGTGAAGCCATAAAAGGTTCTGTCAGTCAAGTGTGTCATACAAATTCCTTTGGC 1620  
QY 1621 ACATCTTGTGAGACGATCCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCAATCCAGCAAC 1680  
DB 1621 ACATCTTGTGAGACGATCCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCAATCCAGCAAC 1680  
QY 1681 ATCTCATTTCTATGCAACAAATGCTGTTGCTCTCTCTCATTTGTCGTTTAAACCTGCTA 1740  
DB 1681 ATCTCATTTCTATGCAACAAATGCTGTTGCTCTCTCTCATTTGTCGTTTAAACCTGCTA 1740  
QY 1741 ATTTGTCACAGTACAAAAGCAAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
DB 1741 ATTTGTCACAGTACAAAAGCAAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
QY 1801 ACOGGCTCTCAGATAATGAGTACTTCTACGTTGATTTTCAAGAAATATGAATATGATCTC 1860  
DB 1801 ACOGGCTCTCAGATAATGAGTACTTCTACGTTGATTTTCAAGAAATATGAATATGATCTC 1860  
QY 1861 AAATGGGAGTTTCCAAAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920  
DB 1861 AAATGGGAGTTTCCAAAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920  
QY 1921 GGAAGAGTATGAAACGCAACAGCTTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
DB 1921 GGAAGAGTATGAAACGCAACAGCTTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
QY 1981 GCGCTCAAAATGCTGAAAAG 2040  
DB 1981 GCGCTCAAAATGCTGAAAAG 2040  
QY 2041 CTCAGATGATCAGCCAGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100  
DB 2041 CTCAGATGATCAGCCAGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100  
QY 2101 ACACTGTGAGGACCAATTTTACTGTTTGTGATTTTGTGATTTGTTGTTGTTGTTGTTGTT 2160  
DB 2101 ACACTGTGAGGACCAATTTTACTGTTTGTGATTTTGTGATTTGTTGTTGTTGTTGTTGTT 2160  
QY 2161 TATCTAAGAGTAAAG 2220  
DB 2161 TATCTAAGAGTAAAG 2220  
QY 2221 AATTTTCAGTTTTCACCACTTTTCCAAATTCACATCCAAATTCAGAGTGTCTGGTTCAAGA 2280  
DB 2221 AATTTTCAGTTTTCACCACTTTTCCAAATTCACATCCAAATTCAGAGTGTCTGGTTCAAGA 2280  
QY 2281 GAAAGTTTCAATACACCGGAGCTCGATCAAAATCTCAGGGCTTTCATGGGAAATTCATTTTCA 2340  
DB 2281 GAAAGTTTCAATACACCGGAGCTCGATCAAAATCTCAGGGCTTTCATGGGAAATTCATTTTCA 2340  
QY 2341 TCTGAAGATGAAATTTGAATATGAAAACCAAAAAGGCTGGAAGAGAGAGAGAGAGAGAGAG 2400  
DB 2341 TCTGAAGATGAAATTTGAATATGAAAACCAAAAAGGCTGGAAGAGAGAGAGAGAGAGAGAG 2400  
QY 2401 GTGCTTACATTTTGAAGATCTTTTGTCTTTCATATCAAGTTTCCCAAGGAAATGGAATTT 2460  
DB 2401 GTGCTTACATTTTGAAGATCTTTTGTCTTTCATATCAAGTTTCCCAAGGAAATGGAATTT 2460  
QY 2461 CTGGAATTTAAGTGTGTGTTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520  
DB 2461 CTGGAATTTAAGTGTGTGTTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520  
QY 2521 GGGAAAGTGGTGAAGATATGCTGTTTGGATTTGGCTCGAGATATCATGAGTATTCACAC 2580

Db	2521		GGGAAAGTGTGGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGATGATTTCAAC	2580
Qy	2581	TATGTTGTCTAGGGCAATGCCGCTCTGCCTGTAAAAATGGATGGCCCCCGAAAGCCCTGTTT	2640	
Db	2581	TATGTTGTCTAGGGCAATGCCGCTCTGCCTGTAAAAATGGATGGCCCCCGAAAGCCCTGTTT	2640	
Qy	2641	GAAGGCATCTACACCATTAAGAGTGATGTCTGGTCAATATGGAAATATCTGTGGGAAATC	2700	
Db	2641	GAAGGCATCTACACCATTAAGAGTGATGTCTGGTCAATATGGAAATATCTGTGGGAAATC	2700	
Qy	2701	TTCTCACTTGGTGTGAATCCTTACCCCTGGCAATCCGGTTGATGCTAACTTCTACAAACTG	2760	
Db	2701	TTCTCACTTGGTGTGAATCCTTACCCCTGGCAATCCGGTTGATGCTAACTTCTACAAACTG	2760	
Qy	2761	ATTTCAAAATGGATTTTAAAAATGGATCAGCCATTTTATGTCTACAGAAATAATCATTTATA	2820	
Db	2761	ATTTCAAAATGGATTTTAAAAATGGATCAGCCATTTTATGTCTACAGAAATAATCATTTATA	2820	
Qy	2821	ATCGAATCCTGCTGGGCTTTTGACTCAAGGAAACGGCCATCTTCCCTAATTTGACTTCG	2880	
Db	2821	ATCGAATCCTGCTGGGCTTTTGACTCAAGGAAACGGCCATCTTCCCTAATTTGACTTCG	2880	
Qy	2881	TTTTTAGGATGTCTCAGCTGGCAGATGCAAGAAAGCGATGTATCAGAAATGTGGATGCCCGT	2940	
Db	2881	TTTTTAGGATGTCTCAGCTGGCAGATGCAAGAAAGCGATGTATCAGAAATGTGGATGCCCGT	2940	
Qy	2941	GTTTGGAAATGTCCTCACCTACCAAAAAGGGGACCTTTTCAGCAGAGAGATGGATTTG	3000	
Db	2941	GTTTGGAAATGTCCTCACCTACCAAAAAGGGGACCTTTTCAGCAGAGAGATGGATTTG	3000	
Qy	3001	GGGCTACTCTCTCCGCAAGGCTCAGGTCGAAGATTCGTAGAGGAAACAATTAGTTTTAAGG	3060	
Db	3001	GGGCTACTCTCTCCGCAAGGCTCAGGTCGAAGATTCGTAGAGGAAACAATTAGTTTTAAGG	3060	
Qy	3061	ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTCACAAAACAAGATTAATTTTCAT	3120	
Db	3061	ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTCACAAAACAAGATTAATTTTCAT	3120	
Qy	3121	CACATAAAGAAAAATCTAATTAACAATCTGCTCCACCAGACTTTTCTCTAGAGAGCCGCTCT	3180	
Db	3121	CACATAAAGAAAAATCTAATTAACAATCTGCTCCACCAGACTTTTCTCTAGAGAGCCGCTCT	3180	
Qy	3181	CGGTTTACTCTGTGTTTCAAAAGGCACTTTGTAAAAATCAAAATCATCTGTCAAGGCAG	3240	
Db	3181	CGGTTTACTCTGTGTTTCAAAAGGCACTTTGTAAAAATCAAAATCATCTGTCAAGGCAG	3240	
Qy	3241	GAGGAGCTGATAATGAACCTTTATTTGGAGCATTGATCTGCATCCAAAGGCCCTCTCAGGCCG	3300	
Db	3241	GAGGAGCTGATAATGAACCTTTATTTGGAGCATTGATCTGCATCCAAAGGCCCTCTCAGGCCG	3300	
Qy	3301	GCTTTGAGTGAATTTGTGTACCTGAAATGACAGTATATCTGTAAATACATAAAAAACAAGC	3360	
Db	3301	GCTTTGAGTGAATTTGTGTACCTGAAATGACAGTATATCTGTGTAAATACATAAAAAACAAGC	3360	
Qy	3361	ATTTTGTCTAAGGAGAGCTAATATGATTTTTTAAAGTCTATGTTTTTAAAAATAATGTAAA	3420	
Db	3361	ATTTTGTCTAAGGAGAGCTAATATGATTTTTTAAAGTCTATGTTTTTAAAAATAATGTAAA	3420	
Qy	3421	TTTTTTCAGCTATTTTAGTGATATATTTTTATGGGTGGGAATAAAATTTCTACTACAGAAAA	3480	
Db	3421	TTTTTTCAGCTATTTTAGTGATATATTTTTATGGGTGGGAATAAAATTTCTACTACAGAAAA	3480	
Qy	3481	AAAAA	3501	
Db	3481	AAAAA	3501	

RESULT 7  
AAQ53503  
ID 'AAQ53503 standard; cDNA; 3501 BP  
XX  
AC AAQ53503;

XX	25-MAR-2003	(updated)	
DT	27-JUN-1994	(first entry)	
XX			
XX	Human flk-2 cDNA.		
XX			
KW	Receptor protein tyrosine kinase; pTK family; foetal liver kinase;		
KW	hflk; primitive; totipotent; haematopoietic cell; stem cell;		
KW	proliferation; stromal cell; ds.		
XX			
OS	Homo sapiens.		
XX			
XX	Key	Location/Qualifiers	
PH	CDS	58..3039	
FT		/*tag= a	
FT		/product= flk-2	
FT	sig_peptide	58..138	
FT		/*tag= b	
FT	mat_peptide	139..3036	
FT		/*tag= c	
XX			
XX	US5270458-A.		
PN			
XX			
XX	14-DEC-1993.		
PD			
XX			
XX	19-NOV-1992;	92US-0977451.	
PF			
XX			
XX	02-APR-1991;	91US-0679666.	
PR			
PR	28-JUN-1991;	91US-0728913.	
PR	15-NOV-1991;	91US-0793065.	
PR	24-DEC-1991;	91US-0813593.	
PR	26-JUN-1992;	92US-0906397.	
PR	12-NOV-1992;	92US-0975049.	
PR	19-NOV-1992;	92US-0977451.	
XX			
XX	(UYPR-) UNIV PRINCETON.		
PA			
XX			
XX	Lemischka IR;		
PI			
XX			
XX	WPI; 1993-405021/50.		
DR			
DR	P-PSDB; AAR44995.		
XX			
XX			
PT	Isolated nucleic acid molecules of hematopoietic stem cell		
PT	receptor flk-2 - encoding mammalian receptor protein tyrosine		
FT	kinases expressed in primitive haematopoietic cells		
FT			
XX			
XX	Disclosure; Fig lb; 60pp; English.		
PS			
XX			
XX	Nucleic acid sequences coding for murine flk-2 and specified		
CC	subfragments of it are claimed. The human flk-2 coding sequence		
CC	(i.e. AA953503) is also disclosed. The flk-2 polypeptides represent		
CC	a new class of receptor protein tyrosine kinases which are		
CC	expressed only in primitive haematopoietic cells.		
CC	(Updated on 25-MAR-2003 to correct PF field.)		
XX			
XX	Sequence 3501 BP; 1068 A; 708 C; 785 G; 940 T; 0 other;		

	Query Match	100.0%;	Score 3499.4;	DB 14;	Length 3501;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 3500;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	1	CGAGGCGGCATCGAGGGCTGGGCGGCGCCCTGGGGGACCCGGGCTCCGGAGGCCATG	60		
Db	1	CGAGGCGGCATCGAGGGCTGGGCGGCGCCCTGGGGGACCCGGGCTCCGGAGGCCATG	60		
Qy	61	CCGCGCTGTGGCGGCGACGCGGGCAACCGTGCCGCTGCTCGTGTGTTTTCTGCAATGATA	120		
Db	61	CCGCGCTGTGGCGGCGACGCGGGCAACCGTGCCGCTGCTCGTGTGTTTTCTGCAATGATA	120		
Qy	121	TTTGGGACTATTCAAAATCAAGATCTTGCCTGTGATCAAGTGTGTTTAAATCAATCAATAAG	180		
Db	121	TTTGGGACTATTCAAAATCAAGATCTTGCCTGTGATCAAGTGTGTTTAAATCAATCAATAAG	180		

181 AACAAATGATTCATCAGTGGGAAGTCATCATCATATCCATGGTATCAGAATCCCCGGAA 240  
181 AACAAATGATTCATCAGTGGGAAGTCATCATCATATCCATGGTATCAGAATCCCCGGAA 240  
241 GACCTCGGGTGTGGTTCAGACCCAGAGCTCAGGACAGTGTACGAAGCTGCCGCTGTG 300  
241 GACCTCGGGTGTGGTTCAGACCCAGAGCTCAGGACAGTGTACGAAGCTGCCGCTGTG 300  
301 GAAGTGGATGATCTGCTTCCATCACA CTGCAAGTCTGCTGCGATGCCCGGAGAACATT 360  
301 GAAGTGGATGATCTGCTTCCATCACA CTGCAAGTCTGCTGCGATGCCCGGAGAACATT 360  
361 TCCTGTCTCGGGTCTTTAAGCACAGCTCCCTGAAATGCCAGCCACATTTTGAATTACAA 420  
361 TCCTGTCTCGGGTCTTTAAGCACAGCTCCCTGAAATGCCAGCCACATTTTGAATTACAA 420  
421 AACAGAGAGTGTTCATCGTCAATTTTGAATAATGACAGAAACCCAGCTGGAGATAC 480  
421 AACAGAGAGTGTTCATCGTCAATTTTGAATAATGACAGAAACCCAGCTGGAGATAC 480  
481 CTACTTTTATTCAGAGTGAAGCTACCAATTTACACNAATTTGTTTACAGTGAATAGA 540  
481 CTACTTTTATTCAGAGTGAAGCTACCAATTTACACNAATTTGTTTACAGTGAATAGA 540  
541 AATACCCCTGCTTTACACAATTAAGAGACCTTTACTTTAGAAAAATGGAACCCAGAGGCC 600  
541 AATACCCCTGCTTTACACAATTAAGAGACCTTTACTTTAGAAAAATGGAACCCAGAGGCC 600  
601 CTGCTCTGCATATCTGAGAGCGTTCCAGAGCGGATCGTGGAAATGGGTGCTTTGGATTCA 660  
601 CTGCTCTGCATATCTGAGAGCGTTCCAGAGCGGATCGTGGAAATGGGTGCTTTGGATTCA 660  
661 CAGGGGAAAGCTGTAAAGAGAAAGTCCAGCTGTTGTTTAAAAAGGAGGAAAAAGTGTCT 720  
661 CAGGGGAAAGCTGTAAAGAGAAAGTCCAGCTGTTGTTTAAAAAGGAGGAAAAAGTGTCT 720  
721 CATGAATTTATTTGGGACGACATAAGTGTCTGCCAGAAATGAACCTGGGACGGAATGC 780  
721 CATGAATTTATTTGGGACGACATAAGTGTCTGCCAGAAATGAACCTGGGACGGAATGC 780  
781 ACCAGGCTGTTCACAATAGATCTAAATCAAACTCCTCAGACCAATTTGCCACAATTTAT 840  
781 ACCAGGCTGTTCACAATAGATCTAAATCAAACTCCTCAGACCAATTTGCCACAATTTAT 840  
841 CTTAAAGTAGGGAAACCCCTTATGGAATAGTGCAGAAAGCTGTTCAATGGAATTC 900  
841 CTTAAAGTAGGGAAACCCCTTATGGAATAGTGCAGAAAGCTGTTCAATGGAATTC 900  
901 GGCTCACTCGGGAAATTAGAAACAAAGCACTCAGAGGGCACTACTTTGAGATGAT 960  
901 GGCTCACTCGGGAAATTAGAAACAAAGCACTCAGAGGGCACTACTTTGAGATGAT 960  
961 ACCTATTCAAACAACAGAACTATGATACGGATCTGTTGCTTTGTTATCATCAGTGGCA 1020  
961 ACCTATTCAAACAACAGAACTATGATACGGATCTGTTGCTTTGTTATCATCAGTGGCA 1020  
1021 AGAAACGACACCGGATACACTGTTTCTCTTTCAAGCACTCCAGTCAATCAGCTTTG 1080  
1021 AGAAACGACACCGGATACACTGTTTCTCTTTCAAGCACTCCAGTCAATCAGCTTTG 1080  
1081 GTTACCACTAGGAAGGATTTATAAATGCTACCAATTTCAAGTGAAGATTTAGAAATT 1140  
1081 GTTACCACTAGGAAGGATTTATAAATGCTACCAATTTCAAGTGAAGATTTAGAAATT 1140  
1141 GACCAATATGAAGTGTGTTTTCTGTCAGGTTTAAAGCTACCCACCAATCAGATGT 1200  
1141 GACCAATATGAAGTGTGTTTTCTGTCAGGTTTAAAGCTACCCACCAATCAGATGT 1200  
1201 ACCTGGAACCTTCTCTCGAAATCATTTCTTGTGAGCAAAAGGGTCTTTGATAACGGATAC 1260  
1201 ACCTGGAACCTTCTCTCGAAATCATTTCTTGTGAGCAAAAGGGTCTTTGATAACGGATAC 1260  
1261 AGCATATCCAAAGTTTTGCAATCATAGCACACCGCAGGAGAAATATATATTCATTCAGAA 1320

1261 AGCATATCCAAAGTTTTGCAATCATAGCACACCGCAGGAGAAATATATATTCATTCAGAA 1320  
1321 AATGATGATGCCCAATTTTACCAAAATGTTTCA CGCTGAATATATAAGAGGAAACCTCAAGTG 1380  
1321 AATGATGATGCCCAATTTTACCAAAATGTTTCA CGCTGAATATATAAGAGGAAACCTCAAGTG 1380  
1381 CTGCGAGAAGCATCGGCAAGTCAAGCGCTCTGTTTCTCGATGGATATCCCATTCATCT 1440  
1381 CTGCGAGAAGCATCGGCAAGTCAAGCGCTCTGTTTCTCGATGGATATCCCATTCATCT 1440  
1441 TGGACCTTGGAAAGAGTGTTCAGACAAGTCTCCCAACTGCA CAGAAGAGATCACAGAAGA 1500  
1441 TGGACCTTGGAAAGAGTGTTCAGACAAGTCTCCCAACTGCA CAGAAGAGATCACAGAAGA 1500  
1501 GTCTGGAATAGAAAGGCTTAACAGAAAAAGTGTGGA CAGTGGGTGTCAGCACTACTCTA 1560  
1501 GTCTGGAATAGAAAGGCTTAACAGAAAAAGTGTGGA CAGTGGGTGTCAGCACTACTCTA 1560  
1561 AACATGATGAGGACCATATAAAGGGTTCCTGTCAGAGTCTGTCATACAAATTCCTCTGGC 1620  
1561 AACATGATGAGGACCATATAAAGGGTTCCTGTCAGAGTCTGTCATACAAATTCCTCTGGC 1620  
1621 ACATCTTGTGAGACGATCCTTTTAAAGTCTCCAGSCCCTTCCCTTTTCATCCACAGCAAC 1680  
1621 ACATCTTGTGAGACGATCCTTTTAAAGTCTCCAGSCCCTTCCCTTTTCATCCACAGCAAC 1680  
1681 ATCTCATTTCTATGCAACAATTTGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740  
1681 ATCTCATTTCTATGCAACAATTTGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740  
1741 ATTTGTCAAGTACAAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGTG 1800  
1741 ATTTGTCAAGTACAAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGTG 1800  
1801 ACCGGCTCTCAGATAATGATGATCTCTACGTTGATTTTTCAGAGATATGATATGATCTC 1860  
1801 ACCGGCTCTCAGATAATGATGATCTCTACGTTGATTTTTCAGAGATATGATATGATCTC 1860  
1861 AAATGGGAGTTTCCAGAGAAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920  
1861 AAATGGGAGTTTCCAGAGAAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920  
1921 GGAAGAGTATGAAACGCAAGCTTTATGGAATTTAGCAAAA CAGGAGTCTCAATCCAGTT 1980  
1921 GGAAGAGTATGAAACGCAAGCTTTATGGAATTTAGCAAAA CAGGAGTCTCAATCCAGTT 1980  
1981 GCGTCAAAATGCTGAAAGAAAAAGCAGACGCTCTGAAAGAGAGGCACTCATGTGAGAA 2040  
1981 GCGTCAAAATGCTGAAAGAAAAAGCAGACGCTCTGAAAGAGAGGCACTCATGTGAGAA 2040  
2041 CTCAAGATGATGACCCAGCTGGGAAGCCACAGAAATTTTGAACCTGCTGGGGGCGTGC 2100  
2041 CTCAAGATGATGACCCAGCTGGGAAGCCACAGAAATTTTGAACCTGCTGGGGGCGTGC 2100  
2101 ACATGTGAGGACCAATTTTACTTTGATTTTGAATACTGTTGCTATGTTGATCTTCTCAAC 2160  
2101 ACATGTGAGGACCAATTTTACTTTGATTTTGAATACTGTTGCTATGTTGATCTTCTCAAC 2160  
2161 TATCTAAGAGTAAAGAGAAAAATTTTCA CAGGACTTTGGA CAGAGATTTTCAAGGAAAC 2220  
2161 TATCTAAGAGTAAAGAGAAAAATTTTCA CAGGACTTTGGA CAGAGATTTTCAAGGAAAC 2220  
2221 AATTTCAAGTGTTCACCCACTTTCCCAATCATCTCCAAATTTCCAGCATGCTGGTTCAGA 2280  
2221 AATTTCAAGTGTTCACCCACTTTCCCAATCATCTCCAAATTTCCAGCATGCTGGTTCAGA 2280  
2281 GAAGTTTCAGATACACCGGACTCGGATCAATCTCAGGGCTTCATGGGAATTCATTTTCA 2340  
2281 GAAGTTTCAGATACACCGGACTCGGATCAATCTCAGGGCTTCATGGGAATTCATTTTCA 2340  
2341 TCTGAAGATGAATTTGAATATGAAAAACCAAAAAGGCTGGAAGAGGAGGACTTGAAT 2400

Db 2341 TCTGAAGATGAATTGAATATGATAAACAACAAAAAGGCTGGAAAGAGGAGGACTTGAAT 2400  
Qy 2401 GTGCTTACATTGAAGATCTTTTGTGCTTTCATATCAAGTTGCCAAAGGAATGGAATTT 2460  
Db 2401 GTGCTTACATTGAAGATCTTTTGTGCTTTCATATCAAGTTGCCAAAGGAATGGAATTT 2460  
Qy 2461 CTGGAATTTAAGTCTGTGTTTCAGAGAGACTGCGCCGAGGAACGCTTGTCAACCCAC 2520  
Db 2461 CTGGAATTTAAGTCTGTGTTTCAGAGAGACTGCGCCGAGGAACGCTTGTCAACCCAC 2520  
Qy 2521 GGGAAAGTGGTGAAGATATGATGCTTTGGATTGCTCGAGATATCATGAGTGATTTCAAC 2580  
Db 2521 GGGAAAGTGGTGAAGATATGATGCTTTGGATTGCTCGAGATATCATGAGTGATTTCAAC 2580  
Qy 2581 TATGTTCTCAGGGGCAATGCCCGTCTGCTGTAAAAATGAGTGGCCCGGAAAGCCTGTTT 2640  
Db 2581 TATGTTCTCAGGGGCAATGCCCGTCTGCTGTAAAAATGAGTGGCCCGGAAAGCCTGTTT 2640  
Qy 2641 GAAGGCATCTACACATTAAGAGTGATGCTGGTGCATATGGAATATTACTGTGGGAATC 2700  
Db 2641 GAAGGCATCTACACATTAAGAGTGATGCTGGTGCATATGGAATATTACTGTGGGAATC 2700  
Qy 2701 TTCTCACTTGTGTGAATCCTTACCTGGCAATTCGGTTGATGCTTAACCTTCAAAAACCTG 2760  
Db 2701 TTCTCACTTGTGTGAATCCTTACCTGGCAATTCGGTTGATGCTTAACCTTCAAAAACCTG 2760  
Qy 2761 ATTCAAAATGGAATTTAAAATGGATCAGCAATTTTATGCTACAGAAAGAAATATACATTATA 2820  
Db 2761 ATTCAAAATGGAATTTAAAATGGATCAGCAATTTTATGCTACAGAAAGAAATATACATTATA 2820  
Qy 2821 ATGCAATCTGCTGGGCTTTTGAATCAAGGAAAGGGCCATCTTCCCTAAATTTGACTTGG 2880  
Db 2821 ATGCAATCTGCTGGGCTTTTGAATCAAGGAAAGGGCCATCTTCCCTAAATTTGACTTGG 2880  
Qy 2881 TTTTGTAGATCTCAGTGGAGATCGAAGAGCGATGATATCAGATGTCGATGGCGGT 2940  
Db 2881 TTTTGTAGATCTCAGTGGAGATCGAAGAGCGATGATATCAGATGTCGATGGCGGT 2940  
Qy 2941 GTTTCGGAATGCTCTCACACCTACCAAAAGGCGACCTTTTTCAGCAGAGAGATGGAATTG 3000  
Db 2941 GTTTCGGAATGCTCTCACACCTACCAAAAGGCGACCTTTTTCAGCAGAGAGATGGAATTG 3000  
Qy 3001 GGGCTACTCTCTCGAGGCTCAGGTGGAAGTTGCTAGAGAAACAATTTAGTTTAAAG 3060  
Db 3001 GGGCTACTCTCTCGAGGCTCAGGTGGAAGTTGCTAGAGAAACAATTTAGTTTAAAG 3060  
Qy 3061 ACTTCATCCCTCCACCTATCCCTAACAGCTGATGATTTACCAACACAGATTAATTTTAT 3120  
Db 3061 ACTTCATCCCTCCACCTATCCCTAACAGCTGATGATTTACCAACACAGATTAATTTTAT 3120  
Qy 3121 CACTAAAAGAAATCTATTATCAACTGCTTTCACAGAGCTTTTCTCTAGAGCGGCTCT 3180  
Db 3121 CACTAAAAGAAATCTATTATCAACTGCTTTCACAGAGCTTTTCTCTAGAGCGGCTCT 3180  
Qy 3181 GCGTTTACTCTGTTTCAAAGGACTTTTGTAAATCAATCATCTCTGTCAACAGGCGAG 3240  
Db 3181 GCGTTTACTCTGTTTCAAAGGACTTTTGTAAATCAATCATCTCTGTCAACAGGCGAG 3240  
Qy 3241 GAGGAGCTGATTAATCAACTTTTATGGAGCATGATCTGCATCCAGGCGCTTCTCAGGCGG 3300  
Db 3241 GAGGAGCTGATTAATCAACTTTTATGGAGCATGATCTGCATCCAGGCGCTTCTCAGGCGG 3300  
Qy 3301 GCTTGAAGTGAATTTGTACCTGAACTCAGATATATTCTTGTAAATACATAAAACAAAGC 3360  
Db 3301 GCTTGAAGTGAATTTGTACCTGAACTCAGATATATTCTTGTAAATACATAAAACAAAGC 3360  
Qy 3361 ATTTTGTAGGAGAGCTAATATGATTTTAAAGTCTATGTTTAAATTAATATGTAAA 3420  
Db 3361 ATTTTGTAGGAGAGCTAATATGATTTTAAAGTCTATGTTTAAATTAATATGTAAA 3420  
Qy 3421 TTTTTCAGCTATTAGTATATATTTTATGGTGGGAATAAATTTCTACTACAGAAAAA 3480  
Db 3421 TTTTTCAGCTATTAGTATATATTTTATGGTGGGAATAAATTTCTACTACAGAAAAA 3480

Qy 3481 AAAAAAAAAAAAAAAAAAAAA 3501  
Db 3481 AAAAAAAAAAAAAAAAAAAAA 3501

## RESULT 8

AAQ40915  
ID AAQ40915 standard; cDNA; 3501 BP.

XX AAQ40915;

XX 25-MAR-2003 (updated)

DT 19-OCT-1993 (first entry)

XX Human flk-2 cDNA.

XX Murine; receptor; protein; tyrosine kinase; pTK; flk-2; primitive;  
KW hematopoietic cell; mature; family; conserved; region;  
KW catalytic domain; c-kit; fetal liver kinase; flk; fetal; spleen;  
KW thymus; adult; brain; bone marrow; multipotential; CFU-Blast colony;  
KW hierarchy; transduction; T-lymphoid; lineage; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 58..3039

XX sig\_peptide /\*tag= a

XX /\*tag= b

XX mat\_peptide /note= "Hydrophobic leader sequence"

XX 139..3036

XX /\*tag= c

XX WO9310136-A1.

XX 27-MAY-1993.

XX 16-NOV-1992; 92WO-US09893.

XX 15-NOV-1991; 91US-0793065.

XX (UYPR-) UNIV PRINCETON.

XX Lemischka IR;

XX WPI; 1993-182479/22.

XX P-PSDB; AAR37503.

XX Totipotent haematopoietic stem cell receptors, their ligands and  
PT DNA sequences - for treating anaemia(s) and bone marrow damage  
PT due to e.g. cancer chemotherapy or radiotherapy

XX Claim 9; Fig 1b; 127pp; English.

XX This sequence encodes the human receptor protein tyrosine kinase  
XX (pTK), flk-2. This nucleic acid is expressed in primitive hemato-  
XX poietic cells and not in mature hematopoietic cells. Members of  
XX this family of pTK's can be recognised by the conserved amino acid  
XX regions in the catalytic domain. This family of pTK's also contains  
XX c-kit. These new receptors are termed fetal liver kinases (flk's)  
XX after the tissue in which they were discovered. flk-2 is also  
XX expressed in fetal spleen, fetal thymus, adult brain and adult  
XX bone marrow. flk-2 is expressed in individual multipotential CFU-  
XX Blast colonies capable of generating numerous multilineage colonies  
XX upon replating. It is likely therefore, that flk-2 is expressed in  
XX the entire primitive portion of the hematopoietic hierarchy. This is  
XX consistent with flk-2 being important in transducing putative self-  
XX renewal signals from the environment. flk-2 is the first receptor  
XX pTK known to be expressed in the T-lymphoid lineage.  
XX (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 3501 BP; 1068 A; 708 C; 785 G; 940 T; 0 other;

Query Match				100.0%; Score 3499.4; DB 14; Length 3501;			
Best Local Similarity				100.0%; Pred No. 0;			
Matches 3500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;							
Qy	1	CGAGCGGCATCCGAGGGCTGGCGCGCCCTGGGGGACCCCGGCTCCGGAGGCCATG	60				
Db	1	CGAGCGGCATCCGAGGGCTGGCGCGCCCTGGGGGACCCCGGCTCCGGAGGCCATG	60				
Qy	61	CGGCGTGGCGCGGACCGCGGACCGTGGCGCTGCTGTTTCTGCAATGATA	120				
Db	61	CGGCGTGGCGCGGACCGCGGACCGTGGCGCTGCTGTTTCTGCAATGATA	120				
Qy	121	TTTGGGACTATTACAAATCAAGATCTGCCCTGTGATCAAGTGTGTTTAAATCAATAG	180				
Db	121	TTTGGGACTATTACAAATCAAGATCTGCCCTGTGATCAAGTGTGTTTAAATCAATAG	180				
Qy	181	AACAATGATTCATAGTGGGGAAGTCATCATCATATCCCATGGTATCAGAATCCCGGAA	240				
Db	181	AACAATGATTCATAGTGGGGAAGTCATCATCATATCCCATGGTATCAGAATCCCGGAA	240				
Qy	241	GACCTCGGCTGTGGTGGAGCCCGAGCTCAGGACAGTGTACGAAGCTCCGCTGTG	300				
Db	241	GACCTCGGCTGTGGTGGAGCCCGAGCTCAGGACAGTGTACGAAGCTCCGCTGTG	300				
Qy	301	GAAATGGGATGATCTGCTTCCATCACCTGCAAGTCTGGTGCATGCCCGGACACATT	360				
Db	301	GAAATGGGATGATCTGCTTCCATCACCTGCAAGTCTGGTGCATGCCCGGACACATT	360				
Qy	361	TCCTGTCTCTGGGTCTTTAAGCACAGCTCCCTGAAATGGCCAGCCACATTTGATTTACAA	420				
Db	361	TCCTGTCTCTGGGTCTTTAAGCACAGCTCCCTGAAATGGCCAGCCACATTTGATTTACAA	420				
Qy	421	AACAGAGAGTGTGTTTCCATGTCATTTTGAATAATGACAGAAACCAAGCTCGAGATAC	480				
Db	421	AACAGAGAGTGTGTTTCCATGTCATTTTGAATAATGACAGAAACCAAGCTCGAGATAC	480				
Qy	481	CTACTTTTATTCAGAGTGAAGTACCAATACAAATATGTTTACAGTGAATATAGA	540				
Db	481	CTACTTTTATTCAGAGTGAAGTACCAATACAAATATGTTTACAGTGAATATAGA	540				
Qy	541	AATACCTCTGCTTTACACATTAAGAGACCTTACTTTAGAAAATGGAACCCAGAGCGC	600				
Db	541	AATACCTCTGCTTTACACATTAAGAGACCTTACTTTAGAAAATGGAACCCAGAGCGC	600				
Qy	601	CTGFTCTCATATCTGAGAGCGTCCAGAGCGATCGTGGATGGTGTGCTTTCGATTC	660				
Db	601	CTGFTCTCATATCTGAGAGCGTCCAGAGCGATCGTGGATGGTGTGCTTTCGATTC	660				
Qy	661	CAGGGGAAAGCTGTAAAGAGAAAGTCCAGCTGTTTAAAGAGGAGAAAGTGCTT	720				
Db	661	CAGGGGAAAGCTGTAAAGAGAAAGTCCAGCTGTTTAAAGAGGAGAAAGTGCTT	720				
Qy	721	CATGAATTTATTTGGACGAGCATAGTGTGTCGAGAAATGAATGAACTGGCGGGAATGC	780				
Db	721	CATGAATTTATTTGGACGAGCATAGTGTGTCGAGAAATGAATGAACTGGCGGGAATGC	780				
Qy	781	ACCAGGCTGTTTCAATAGATCTAAATCAAACTCTCTCAGACCAATTTGCCAATTTAT	840				
Db	781	ACCAGGCTGTTTCAATAGATCTAAATCAAACTCTCTCAGACCAATTTGCCAATTTAT	840				
Qy	841	CTTAAAGTGGGAAACCTTTATGGATTAAGTGCAGAAAGCTGTGTAACCATGGATTC	900				
Db	841	CTTAAAGTGGGAAACCTTTATGGATTAAGTGCAGAAAGCTGTGTAACCATGGATTC	900				
Qy	901	GGGCTCACCTGGGATTAAGAAACCAAGCTCGAGGAGGCACTTCTTGGATGAT	960				
Db	901	GGGCTCACCTGGGATTAAGAAACCAAGCTCGAGGAGGCACTTCTTGGATGAT	960				
Qy	961	ACCTATTCAAACAAACAGAACTATGATCGGATTTCTGTTTGTATCATCATGAGCA	1020				
Db	961	ACCTATTCAAACAAACAGAACTATGATCGGATTTCTGTTTGTATCATCATGAGCA	1020				

Qy	1021	AGAAACGACACCGGATCTACACTTGTCTCTTCAAAGCATCCAGTCAATCAGCTTGG	1080				
Db	1021	AGAAACGACACCGGATCTACACTTGTCTCTTCAAAGCATCCAGTCAATCAGCTTGG	1080				
Qy	1081	GTTCACATCGTAGGAAAGGGATTTATAAATGCTACCAATTCAGTGAAGATTTATGAAAT	1140				
Db	1081	GTTCACATCGTAGGAAAGGGATTTATAAATGCTACCAATTCAGTGAAGATTTATGAAAT	1140				
Qy	1141	GACCAATATGAAGAGTGTGTTTCTGTGAGTTTAAAGCTTACCCACAAATCAGATGT	1200				
Db	1141	GACCAATATGAAGAGTGTGTTTCTGTGAGTTTAAAGCTTACCCACAAATCAGATGT	1200				
Qy	1201	ACGTGACCTTCTCTCGAAATCATTTCTCTGTGAGCAAAAGGCTTCTGATACGGATAC	1260				
Db	1201	ACGTGACCTTCTCTCGAAATCATTTCTCTGTGAGCAAAAGGCTTCTGATACGGATAC	1260				
Qy	1261	AGCATATCAAAAGTGTTCATATCAAGCACAGCCAGGAGATATATATTTCCATGCAGAA	1320				
Db	1261	AGCATATCAAAAGTGTTCATATCAAGCACAGCCAGGAGATATATATTTCCATGCAGAA	1320				
Qy	1321	AATGATGATGCCAATTTTACAAATGTTTACGCTGAATATAAGAGGAAACCTCAAGTG	1380				
Db	1321	AATGATGATGCCAATTTTACCAAAATGTTTACGCTGAATATAAGAGGAAACCTCAAGTG	1380				
Qy	1381	CTCGCAGAGCATCGGCAAGTCAGCGCTCTGTTTCTCGGATGATACCCATTTCCATCT	1440				
Db	1381	CTCGCAGAGCATCGGCAAGTCAGCGCTCTGTTTCTCGGATGATACCCATTTCCATCT	1440				
Qy	1441	TGGACCTCGAAGAGTGTTCAGACAAAGTCTCCCAACTGCACAGAGAGATCACAGAGGA	1500				
Db	1441	TGGACCTCGAAGAGTGTTCAGACAAAGTCTCCCAACTGCACAGAGAGATCACAGAGGA	1500				
Qy	1501	GTCTGGAATAGAAAGGCTTACAGAAAAGTGTGTCAGAGTGGTGTGTCAGAGTACTCTA	1560				
Db	1501	GTCTGGAATAGAAAGGCTTACAGAAAAGTGTGTCAGAGTGGTGTGTCAGAGTACTCTA	1560				
Qy	1561	AACATGAGTGAAGCCATAAAAGGGTCTGTCAGAGTGGTGTGTCATTAATTCCTTGGC	1620				
Db	1561	AACATGAGTGAAGCCATAAAAGGGTCTGTCAGAGTGGTGTGTCATTAATTCCTTGGC	1620				
Qy	1621	ACATCTGTGAGAGCATCTTTTAACTCTCCAGGCCCCCTTCCCTTTTATCCAGACAC	1680				
Db	1621	ACATCTGTGAGAGCATCTTTTAACTCTCCAGGCCCCCTTCCCTTTTATCCAGACAC	1680				
Qy	1681	ATCTCATTTCTATGCAAACTTGGTGTGTCCTCTCTTCTTCTGTTTAAACCTGCTA	1740				
Db	1681	ATCTCATTTCTATGCAAACTTGGTGTGTCCTCTCTTCTCTTCTGTTTAAACCTGCTA	1740				
Qy	1741	ATTTGTCAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG	1800				
Db	1741	ATTTGTCAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG	1800				
Qy	1801	ACCGCTCTCAGATATGATGATCTTCTACGTTGATTTTTCAGAGATATGATATGATCTC	1860				
Db	1801	ACCGCTCTCAGATATGATGATCTTCTACGTTGATTTTTCAGAGATATGATATGATCTC	1860				
Qy	1861	AAATGGGAGTTCCTCAAGAGAAAATTTAGAGTGTGGAGAGTCTAGGATCAGGTCTTTT	1920				
Db	1861	AAATGGGAGTTCCTCAAGAGAAAATTTAGAGTGTGGAGAGTCTAGGATCAGGTCTTTT	1920				
Qy	1921	GGAAAGTGTGAAACGCAAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGTT	1980				
Db	1921	GGAAAGTGTGAAACGCAAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGTT	1980				
Qy	1981	GCCGTCAAAATGCTGAAGAAAAGCAGCTCTGAAAGAGGAGCTCATGTTCAGAA	2040				
Db	1981	GCCGTCAAAATGCTGAAGAAAAGCAGCTCTGAAAGAGGAGCTCATGTTCAGAA	2040				
Qy	2041	CTCAAGTGTGAGCCAGCTGGGAGGCCACAGAGATTTGTGAACTCTCGGGGCGTGC	2100				
Db	2041	CTCAAGTGTGAGCCAGCTGGGAGGCCACAGAGATTTGTGAACTCTCGGGGCGTGC	2100				
Qy	2101	ACATGTGAGGACCAATTTTCTTGAATTTTGAATATCTGTTGTTGTTGTTCTTCAAC	2160				





CC to be uncommitted. Therefore, thymocytes expressing flk-2 may be  
CC multipotential. flk-2 is the first receptor tyrosine kinase known to  
CC be expressed in the T-lymphoid lineage.  
CC (Updated on 25-MAR-2003 to correct PN field.)

SQ Sequence 3501 BP; 1067 A; 708 C; 786 G; 940 T; 0 other;

Query Match	99.9%	Score 3497.8;	DB 14;	Length 3501;
Best Local Similarity	99.9%	Pred. No. 0;		
Matches 3499;	Conservative	0;	Mismatches 2;	Indels 0; Gaps 0;

QY	1	CGAGGCGGCATCCGAGGGCTGGCCGGCGCCCTGGGGACCCCGGGCTCCGGAGGCCATG	60
Db	1	CGAGGCGGCATCCGAGGGCTGGCCGGCGCCCTGGGGACCCCGGGCTCCGGAGGCCATG	60
QY	61	CGGGCGTTGGCGGCGAGCGCGGGACCTGGCGCGTCTGTTGTTTTTCTGCAATGATA	120
Db	61	CGGGCGTTGGCGGCGGACGCGGGCACCGTGCGCGTCTGTTGTTTTTCTGCAATGATA	120
QY	121	TTTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTTAATCAATCATAG	180
Db	121	TTTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTTAATCAATCATAG	180
QY	181	AACAAATGATTCATGATGGGGAAGTCATCATATCCCATGGTATCAGAATCCCGGAA	240
Db	181	AACAAATGATTCATGATGGGGAAGTCATCATATCCCATGGTATCAGAATCCCGGAA	240
QY	241	GA CCTCGGGTGTGCGTTGAGACCCAGAGCTCAGGACAGGTGACGAAGCTGCCGTGTG	300
Db	241	GA CCTCGGGTGTGCGTTGAGACCCAGAGCTCAGGACAGGTGACGAAGCTGCCGTGTG	300
QY	301	GAAGTGGATGTATCTGCTTCCATCACACTGCAAGTGTGTCGATGCGCCAGGGAACAT	360
Db	301	GAAGTGGATGTATCTGCTTCCATCACACTGCAAGTGTGTCGATGCGCCAGGGAACAT	360
QY	361	TCTCTCTCTGGGTCTTTAAGCACAGCTGCCGAATGGCCAGGCACATTTTGATTTACAA	420
Db	361	TCTCTCTCTGGGTCTTTAAGCACAGCTGCCGAATGGCCAGGCACATTTTGATTTACAA	420
QY	421	AACAGAGAGTGTGTTTCCATGGTCATTTGAAATGACAGAAACCCCAAGCTGGAGATAC	480
Db	421	AACAGAGAGTGTGTTTCCATGGTCATTTGAAATGACAGAAACCCCAAGCTGGAGATAC	480
QY	481	CTACTTTTATTCAGAGTGAAGCTACCAATTTACAAATATGTTTACAGTGAATAGA	540
Db	481	CTACTTTTATTCAGAGTGAAGCTACCAATTTACAAATATGTTTACAGTGAATAGA	540
QY	541	AATACCTCTGCTTTACACATTAAGAGACCTTACTTTAGAAAATGAAAACCAAGACGC	600
Db	541	AATACCTCTGCTTTACACATTAAGAGACCTTACTTTAGAAAATGAAAACCAAGACGC	600
QY	601	CTGCTGTCATATCTGAGAGCGTTCCAGAGCCGATCGTGAATGGTGTCTTGGCATTC	660
Db	601	CTGCTGTCATATCTGAGAGCGTTCCAGAGCCGATCGTGAATGGTGTCTTGGCATTC	660
QY	661	CAGGGGGAAGCTGTAAGAGAAAGTGCAGTGTGTTTAAAAAGGAGAAAAGTGCTT	720
Db	661	CAGGGGGAAGCTGTAAGAGAAAGTGCAGTGTGTTTAAAAAGGAGAAAAGTGCTT	720
QY	721	CATGAATTTATTTGGGACGACATAGTGTGTCAGGAATGAACTGGGCGAGGGAATGC	780
Db	721	CATGAATTTATTTGGGACGACATAGTGTGTCAGGAATGAACTGGGCGAGGGAATGC	780
QY	781	ACCAGGCTGTTACAAATAGATCTAAATCAAACCTCCACAGCACATGTCACAAATATT	840
Db	781	ACCAGGCTGTTACAAATAGATCTAAATCAAACCTCCACAGCACATGTCACAAATATT	840
QY	841	CTTAAAGTAGGGGAACCCCTATGGAATAGGTGCAAAAGCTGTTCA TGTGAACCATGATTC	900
Db	841	CTTAAAGTAGGGGAACCCCTATGGAATAGGTGCAAAAGCTGTTCA TGTGAACCATGATTC	900
QY	901	GGGCTCACCTGGGAATTTAGAAAAAAGACACTCGAGGAGGGCACTACTTTGAGATGAT	960
Db	901	GGGCTCACCTGGGAATTTAGAAAAAAGACACTCGAGGAGGGCACTACTTTGAGATGAT	960

Db	901	GGGCTCACTCGGGAATTGAAAAA	CAAAAGCACTCGAGGAGGGCAAC	TACTACTTTGAGATGAGT	960
Qy	961	ACCTATTCAACAAAACAGAACTAT	GATACGGATTCTGTTTCTTTGTTG	TATCATCATCAGTGGCA	1020
Db	961	ACCTATTCAACAAAACAGAACTAT	GATACGGATTCTGTTTCTTTGTTG	TATCATCATCAGTGGCA	1020
Qy	1021	AGAAACGACACCGGATACCTAC	CTTGTGTTCTCTTCAAAGCATCC	CAGTCAATCAGCTTTG	1080
Db	1021	AGAAACGACACCGGATACCTAC	CTTGTGTTCTCTTCAAAGCATCC	CAGTCAATCAGCTTTG	1080
Qy	1081	GTTTACCACTGCTAGGAAGGGAT	TATTAATGCTTACCATTCAAGTGA	AGATTAATGAAATT	1140
Db	1081	GTTTACCACTGCTAGGAAGGGAT	TATTAATGCTTACCATTCAAGTGA	AGATTAATGAAATT	1140
Qy	1141	GACCAATATGAAGAGTTTGTGTT	CTCTGTCAGGGTTTAAAGCCTACC	CACAAAATCAGATGT	1200
Db	1141	GACCAATATGAAGAGTTTGTGTT	CTCTGTCAGGGTTTAAAGCCTACC	CACAAAATCAGATGT	1200
Qy	1201	ACGTGGACCTTCTCTCGAAATCA	TTTCTCTGTCAGGGTTTAAAGG	GTCTTGATAACGGATC	1260
Db	1201	ACGTGGACCTTCTCTCGAAATCA	TTTCTCTGTCAGGGTTTAAAGG	GTCTTGATAACGGATC	1260
Qy	1261	AGCATATCCCAAGTTTTCGAAAT	CTATTAAGCACCACGAGAGAA	TATATATTCATCGAGAA	1320
Db	1261	AGCATATCCCAAGTTTTCGAAAT	CTATTAAGCACCACGAGAGAA	TATATATTCATCGAGAA	1320
Qy	1321	AATGATGATGCCCAATTTTACAAA	ATGTTTCCAGCTCGAATATAGA	AGGAAACCTCAAGTG	1380
Db	1321	AATGATGATGCCCAATTTTACAAA	ATGTTTCCAGCTCGAATATAGA	AGGAAACCTCAAGTG	1380
Qy	1381	CTCGCAAGACATCGGCAAGTCT	CGAGCTCTGTTTCTCGGATGGA	TACCCATTTACCATCT	1440
Db	1381	CTCGCAAGACATCGGCAAGTCT	CGAGCTCTGTTTCTCGGATGGA	TACCCATTTACCATCT	1440
Qy	1441	TGGACCTTGGAGAGAGTTTTCAGA	CAAGTCTCCCAA	CTGCA	1500
Db	1441	TGGACCTTGGAGAGAGTTTTCAGA	CAAGTCTCCCAA	CTGCA	1500
Qy	1501	GTCCTGGAATAGAAGGCTTACAGA	AAAGTGTGTTGGACAGTGGGT	GTGCGAGCAGTACTCTA	1560
Db	1501	GTCCTGGAATAGAAGGCTTACAGA	AAAGTGTGTTGGACAGTGGGT	GTGCGAGCAGTACTCTA	1560
Qy	1561	AACATGAGTGAAGCCATATAAAG	GGTTTCTTGCTCAAGTGTGTG	CA	1620
Db	1561	AACATGAGTGAAGCCATATAAAG	GGTTTCTTGCTCAAGTGTGTG	CA	1620
Qy	1621	ACATCTTGTGAGACGATCTTTTAA	ACTCTCAGGCCCCCTTCCCTTT	TCATCAAGCAAC	1680
Db	1621	ACATCTTGTGAGACGATCTTTTAA	ACTCTCAGGCCCCCTTCCCTTT	TCATCAAGCAAC	1680
Qy	1681	ATCTCATCTATGCAACAATTTAG	GTGTTGTTGTTGTTGTTGTTG	TTTAA	1740
Db	1681	ATCTCATCTATGCAACAATTTAG	GTGTTGTTGTTGTTGTTGTTG	TTTAA	1740
Qy	1741	ATTGTTGCAAGTACAAAAGCAAT	TTTAGGTATGAAAGCCAGCTAC	AGATGGTACAGGTG	1800
Db	1741	ATTGTTGCAAGTACAAAAGCAAT	TTTAGGTATGAAAGCCAGCTAC	AGATGGTACAGGTG	1800
Qy	1801	ACGGCTCTCTGATTAATGAGTACT	TTCAGTGTGTTTACAGGAAAT	TGAATATGATCTC	1860
Db	1801	ACGGCTCTCTGATTAATGAGTACT	TTCAGTGTGTTTACAGGAAAT	TGAATATGATCTC	1860
Qy	1861	AAATGGGAGTTTCCAAAGCAAA	TTTAGGTGTTTGGGAGGTTACT	AGGATCAGTGTCTTTT	1920
Db	1861	AAATGGGAGTTTCCAAAGCAAA	TTTAGGTGTTTGGGAGGTTACT	AGGATCAGTGTCTTTT	1920
Qy	1921	GGAAAAAGTGAACGCAACAGCT	TATGGAATTTAGCAAAA	CAGGAGTCTCAATCCAGTT	1980
Db	1921	GGAAAAAGTGAACGCAACAGCT	TATGGAATTTAGCAAAA	CAGGAGTCTCAATCCAGTT	1980
Qy	1981	GGCGTCAAAATGCTGAAAGAAA	AAAGCAGACAGCTCTGAAAG	AGAGGCACTCATGTCA	2040
Db	1981	GGCGTCAAAATGCTGAAAGAAA	AAAGCAGACAGCTCTGAAAG	AGAGGCACTCATGTCA	2040





XX PS Claim 8; Fig 1; 66pp; English.

XX CC The STK-1 gene encodes a receptor PTK which is expressed in

XX CC proliferating hematopoietic stem cells but not in quiescent stem

XX CC cells. The STK-1 gene is also expressed in certain malignant cells

XX CC of non-hematopoietic origin. An antisense oligo specific for STK-1

XX CC is an oligo having a sequence (i) capable of forming a stable

XX CC triplex with a portion of the STK-1 gene, or (ii) capable of forming

XX CC a stable duplex with a portion of an mRNA transcript of the STK-1

XX CC gene. Antisense oligos capable of forming a stable duplex with a

XX CC portion of a STK-1 mRNA transcript are given in AA091536 PT and in

XX CC AA091537 and AA091538. The antisense oligos of the invention are useful

XX CC in the treatment of hematologic malignancies characterised by

XX CC STK-1 expression.

XX CC (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 3476 BP; 1044 A; 709 C; 783 G; 940 T; 0 other;

Query Match 99.2%; Score 3474.4; DB 16; Length 3476;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3475; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGAGCGGCATCCGAGGGCTGGCGCGCCCTGGGGACCCCGGGCTCCGGAGGCCATG 60

DB 1 CGAGCGGCATCCGAGGGCTGGCGCGCCCTGGGGACCCCGGGCTCCGGAGGCCATG 60

QY 61 CCGGCGTTGGCGCGAGCGCGGCGACCGTGCCTGGTGTGTTTTCGCAATGATA 120

DB 61 CCGGCGTTGGCGCGAGCGCGGCGACCGTGCCTGGTGTGTTTTCGCAATGATA 120

QY 121 TTTCGGACTATTACAAATCAGATCTGCTGTGATCAAGTGTGTTTAAATCAATAG 180

DB 121 TTTCGGACTATTACAAATCAGATCTGCTGTGATCAAGTGTGTTTAAATCAATAG 180

QY 181 AACAAATGATTCATCAGTGGGGAAGTCATCATCATATCCATGATCAGAAATCCCGGAA 240

DB 181 AACAAATGATTCATCAGTGGGGAAGTCATCATCATATCCATGATCAGAAATCCCGGAA 240

QY 241 GACCTCGGGTGTGGTGGAGTCCAGAGTCCAGGACAGTGTACGAAAGTCCCGCTGTG 300

DB 241 GACCTCGGGTGTGGTGGAGTCCAGAGTCCAGGACAGTGTACGAAAGTCCCGCTGTG 300

QY 301 GAAGTGGATGTATCTGCTTCCATCATCATCTGATCAAGTGTGCTGATGCGCCAGGGAACATT 360

DB 301 GAAGTGGATGTATCTGCTTCCATCATCATCTGATCAAGTGTGCTGATGCGCCAGGGAACATT 360

QY 361 TCCTGTCTCTGGTCTTTAAGCACAGCTCCCTGAATTCGACGCCACATTTTGAATTTACA 420

DB 361 TCCTGTCTCTGGTCTTTAAGCACAGCTCCCTGAATTCGACGCCACATTTTGAATTTACA 420

QY 421 AACAGAGAGTGTTCATGGTCAATTTTGAATAATGACAGAAACCCAGCTGGAGAAATAC 480

DB 421 AACAGAGAGTGTTCATGGTCAATTTTGAATAATGACAGAAACCCAGCTGGAGAAATAC 480

QY 481 CTACTTTTATTCAGAGTGAAGTACCAATTTACAAATATTTGTTTACAGTGAATAGA 540

DB 481 CTACTTTTATTCAGAGTGAAGTACCAATTTACAAATATTTGTTTACAGTGAATAGA 540

QY 541 AATAACCTGCTTTTACAAATTAAGAGAGCTTTTACATTTTGAATAATGGAACCCAGAGGCC 600

DB 541 AATAACCTGCTTTTACAAATTAAGAGAGCTTTTACATTTTGAATAATGGAACCCAGAGGCC 600

QY 601 CTGCTGTGCATATCTGAGAGGTTCCAGAGCCGATCGTGAATGGTGTCTTTCGATTTCA 660

DB 601 CTGCTGTGCATATCTGAGAGGTTCCAGAGCCGATCGTGAATGGTGTCTTTCGATTTCA 660

QY 661 CAGGGGAAAGCTGTAAGAAGAAAGTTCAGCTGTGTTTAAAGAGGAAAGTGTCTT 720

DB 661 CAGGGGAAAGCTGTAAGAAGAAAGTTCAGCTGTGTTTAAAGAGGAAAGTGTCTT 720

QY 721 CATGAATTTTGGGACGACATAAGTGTCTGCCAGTAATGAATGAGTGGGAGGGAATGC 780

DB 721 CATGAATTTTGGGACGACATAAGTGTCTGTGCCAGAAATGAATGCGCAGGGAATGC 780

QY 781 ACCAGGCTGTTCAATAGATCTAAATCAAACTCTCTCAGACCAATTTGCGCACAATTTATTT 840

DB 781 ACCAGGCTGTTCAATAGATCTAAATCAAACTCTCTCAGACCAATTTGCGCACAATTTATTT 840

QY 841 CTTAAAGTAGGGGAACCCCTTATGGAATAGGTGCAAGCTGTTTCTATGTGAACCAATGGAATTC 900

DB 841 CTTAAAGTAGGGGAACCCCTTATGGAATAGGTGCAAGCTGTTTCTATGTGAACCAATGGAATTC 900

QY 901 GGGCTCACCTGGGAATTTAGAAAACAAAGCACTCGAGAGGGCAAACTACTTTTGAGATGAGT 960

DB 901 GGGCTCACCTGGGAATTTAGAAAACAAAGCACTCGAGAGGGCAAACTACTTTTGAGATGAGT 960

QY 961 ACCTATTCAACAAACAGAACTATGATACGGAATCTCTGTTGCTTTTGTATCATCAGTGGCA 1020

DB 961 ACCTATTCAACAAACAGAACTATGATACGGAATCTCTGTTGCTTTTGTATCATCAGTGGCA 1020

QY 1021 AGAAACGACACCGGATACACTTGTCTCTTCAAGCATCCCAAGTCCCAAGTTCAGCTTTG 1080

DB 1021 AGAAACGACACCGGATACACTTGTCTCTTCAAGCATCCCAAGTCCCAAGTTCAGCTTTG 1080

QY 1081 GTTACCATCTGTAGGAAGGATTTTAAATGCTTACCAATTTCAAGTGAAGATTTAGAAAT 1140

DB 1081 GTTACCATCTGTAGGAAGGATTTTAAATGCTTACCAATTTCAAGTGAAGATTTAGAAAT 1140

QY 1141 GACCAATATGAGAGTGTGTTTGTCTGTGAGCAAAAGGGTCTTGATAACGATATC 1200

DB 1141 GACCAATATGAGAGTGTGTTTGTCTGTGAGCAAAAGGGTCTTGATAACGATATC 1200

QY 1201 AGTGGACCTTCTCTCGAAAATCAATTCCTGTGAGCAAAAGGGTCTTGATAACGATATC 1260

DB 1201 AGTGGACCTTCTCTCGAAAATCAATTCCTGTGAGCAAAAGGGTCTTGATAACGATATC 1260

QY 1261 AGCATATCCAAAGTGTGCAATCATTAAGCAACGAGGAGAAATATATATTTCCATGAGAA 1320

DB 1261 AGCATATCCAAAGTGTGCAATCATTAAGCAACGAGGAGAAATATATATTTCCATGAGAA 1320

QY 1321 AATGATGATGCCCAATTTTACCAAAATGTTTACGCTGAATATTAAGAGGAAACCTCAAGTG 1380

DB 1321 AATGATGATGCCCAATTTTACCAAAATGTTTACGCTGAATATTAAGAGGAAACCTCAAGTG 1380

QY 1381 CTGCGAGAAAGCATCGGCAAGTCAAGCTCTCCCACTGCAAGAGATCAAGATATCAAGTATC 1440

DB 1381 CTGCGAGAAAGCATCGGCAAGTCAAGCTCTCCCACTGCAAGAGATCAAGATATCAAGTATC 1440

QY 1441 TGGACCTGGAAGAGTGTTCAGCAAGTCTCCCACTGCAAGAGATCAAGAGAA 1500

DB 1441 TGGACCTGGAAGAGTGTTCAGCAAGTCTCCCACTGCAAGAGATCAAGAGAA 1500

QY 1501 GTCTGGAATGAAAGGCTTAACAGAAAAGTGTGAGCAAGTGTGAGCAAGTGTGAGCAAGTGTG 1560

DB 1501 GTCTGGAATGAAAGGCTTAACAGAAAAGTGTGAGCAAGTGTGAGCAAGTGTGAGCAAGTGTG 1560

QY 1561 AACATGATGAGGCAATTAAGAGGTTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1620

DB 1561 AACATGATGAGGCAATTAAGAGGTTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1620

QY 1621 ACATCTGTGAGAGTCTTTTAACTCTCCAGGCGCTTCCCTTTCATCCAGAGCAAC 1680

DB 1621 ACATCTGTGAGAGTCTTTTAACTCTCCAGGCGCTTCCCTTTCATCCAGAGCAAC 1680

QY 1681 ATCTCAATCTATGCAAACTTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740

DB 1681 ATCTCAATCTATGCAAACTTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740

QY 1741 ATTTGTCAAGTACAAAAAGCAATTTAGTATTAAGAGGCAAGTACAGATGATGATGATGATGATG 1800

DB 1741 ATTTGTCAAGTACAAAAAGCAATTTAGTATTAAGAGGCAAGTACAGATGATGATGATGATGATG 1800

QY 1801 ACCGGCTCTCAGATATGATGATCTTCTAGCTTGAATTTTCAAGAGAAATGAAATATGATATCTC 1860

DB 1801 ACCGGCTCTCAGATATGATGATCTTCTAGCTTGAATTTTCAAGAGAAATGAAATATGATATCTC 1860

```
Qy 1861 AAATGGGAGTTTCCAGAGAAAAATTTAGAGTTTGGGAAAGGTACTAGGATCAGGTGCTTTT 1920
Db 1861 AAATGGGAGTTTCCAGAGAAAAATTTAGAGTTTGGGAAAGGTACTAGGATCAGGTGCTTTT 1920
Qy 1921 GGAAGAATGATGAACGCAACAGCTTATGGAATTTAGCAAAAAGGAGTCTCAATCCAGGTT 1980
Db 1921 GGAAGAATGATGAACGCAACAGCTTATGGAATTTAGCAAAAAGGAGTCTCAATCCAGGTT 1980
Qy 1981 GCCGTCAAAATGCTGAAGAAAAAGCAGACAGCTCTGAAGAGAGGCACTCATGTCAAAA 2040
Db 1981 GCCGTCAAAATGCTGAAGAAAAAGCAGACAGCTCTGAAGAGAGGCACTCATGTCAAAA 2040
Qy 2041 CTCAAGATGATGACCCAGCTGGGAGCCACAGAGATATTTGTAACCTCTGGGGCGGTGC 2100
Db 2041 CTCAAGATGATGACCCAGCTGGGAGCCACAGAGATATTTGTAACCTCTGGGGCGGTGC 2100
Qy 2101 ACATGTCAGACCAATTTACTTGAATTTTGAATCTGTTGATGATGATCTTCTCAAC 2160
Db 2101 ACATGTCAGACCAATTTACTTGAATTTTGAATCTGTTGATGATGATCTTCTCAAC 2160
Qy 2161 TATCTAAGAGTAAAGAGAAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAAACAC 2220
Db 2161 TATCTAAGAGTAAAGAGAAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAAACAC 2220
Qy 2221 AATTTTCAGTTTTTACCCCACTTTTCCAATTCACATCCAAATTCAGCATGCTGTTTCAAGA 2280
Db 2221 AATTTTCAGTTTTTACCCCACTTTTCCAATTCACATCCAAATTCAGCATGCTGTTTCAAGA 2280
Qy 2281 GAAGTTCAGATACACCCGAGCTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCAAC 2340
Db 2281 GAAGTTCAGATACACCCGAGCTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCAAC 2340
Qy 2341 TCTGAAGATGAATTTGAATATGAATATGAATATGAATATGAATATGAATATGAATATGAAT 2400
Db 2341 TCTGAAGATGAATTTGAATATGAATATGAATATGAATATGAATATGAATATGAATATGAAT 2400
Qy 2401 GTGCTTACATTTGAAGATCTTCTTTGCTTTTGCATATCAAGTTTGCACAAAGGAATGGAATTT 2460
Db 2401 GTGCTTACATTTGAAGATCTTCTTTGCTTTTGCATATCAAGTTTGCACAAAGGAATGGAATTT 2460
Qy 2461 CTGGAATTTAAGTCTGTGTTTCAAGAGACCTGGCCGCGCAGGAACGTGCTGTCAACCAC 2520
Db 2461 CTGGAATTTAAGTCTGTGTTTCAAGAGACCTGGCCGCGCAGGAACGTGCTGTCAACCAC 2520
Qy 2521 GGGAAAGTGTGAAGATATGACTTTTGGATTTGGCTCGAGATATCATGATGATTTCCAAAC 2580
Db 2521 GGGAAAGTGTGAAGATATGACTTTTGGATTTGGCTCGAGATATCATGATGATTTCCAAAC 2580
Qy 2581 TATGTTGTGAGGGCAATGCCCGTCTGCTGTAAATGAGTGGCCCGCCGAAAGCCTGTTT 2640
Db 2581 TATGTTGTGAGGGCAATGCCCGTCTGCTGTAAATGAGTGGCCCGCCGAAAGCCTGTTT 2640
Qy 2641 GAAGCATCTACACCATTAAGAGTATGCTGCTGATATGGAATATTAATCTGTGGGAATC 2700
Db 2641 GAAGCATCTACACCATTAAGAGTATGCTGCTGATATGGAATATTAATCTGTGGGAATC 2700
Qy 2701 TTCTCAGTTGTTGTAATCTTACCTGCAATTCGGTTTGGATGCTAACTTCTACAACTG 2760
Db 2701 TTCTCAGTTGTTGTAATCTTACCTGCAATTCGGTTTGGATGCTAACTTCTACAACTG 2760
Qy 2761 ATTCAAAATGGAATTTAAATGGAATCAGCCATTTTATGCTTACAGAAAGAAATATACATTATA 2820
Db 2761 ATTCAAAATGGAATTTAAATGGAATCAGCCATTTTATGCTTACAGAAAGAAATATACATTATA 2820
Qy 2821 ATGCAATCTGCTGGGCTTTTGAATCAAGGAAACCGGCCATCTTCCCTAAATTTGACTTCG 2880
Db 2821 ATGCAATCTGCTGGGCTTTTGAATCAAGGAAACCGGCCATCTTCCCTAAATTTGACTTCG 2880
Qy 2881 TTTTGTAGATGTCAGCTGGCAGTGCAGAGAGCGATGCTATCAGATGCTGATGCGCGT 2940
Db 2881 TTTTGTAGATGTCAGCTGGCAGTGCAGAGAGCGATGCTATCAGATGCTGATGCGCGT 2940
```

```
Qy 2941 GTTTCGGAATGCTCTCAGACCTACCAAAAACAGGCGACCTTTTCAGCAGAGATGGAATTTG 3000
Db 2941 GTTTCGGAATGCTCTCAGACCTACCAAAAACAGGCGACCTTTTCAGCAGAGATGGAATTTG 3000
Qy 3001 GGGCTACTCTCTCCGAGGCTCAGGTGCAAGATTCGTAGAGGAAACAAATTTAGTTTAAAGG 3060
Db 3001 GGGCTACTCTCTCCGAGGCTCAGGTGCAAGATTCGTAGAGGAAACAAATTTAGTTTAAAGG 3060
Qy 3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTTACCAAAAACAAAGATTAATTTTCA 3120
Db 3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTTACCAAAAACAAAGATTAATTTTCA 3120
Qy 3121 CACTAAAAGAAAATCTATATCAACTGCTCTCACAGACTTTTCTCTAGAGCCGTCT 3180
Db 3121 CACTAAAAGAAAATCTATATCAACTGCTCTCACAGACTTTTCTCTAGAGCCGTCT 3180
Qy 3181 GCGTTTACTCTGTTTTCAAAGGCACTTTTGTAAAATCAAAATCATCTCTCACAGGCGAG 3240
Db 3181 GCGTTTACTCTGTTTTCAAAGGCACTTTTGTAAAATCAAAATCATCTCTCACAGGCGAG 3240
Qy 3241 GAGGAGCTGATAATGAACCTTTTATTTGGAGCACTTGTCTGCATCCAGGCGCTTCTCAGGCGG 3300
Db 3241 GAGGAGCTGATAATGAACCTTTTATTTGGAGCACTTGTCTGCATCCAGGCGCTTCTCAGGCGG 3300
Qy 3301 GCITGAGTGAATCTGTACTGCTGAAAGTACAGTATATCTTTGTAATAATACATAAAAACAAAAGC 3360
Db 3301 GCITGAGTGAATCTGTACTGCTGAAAGTACAGTATATCTTTGTAATAATACATAAAAACAAAAGC 3360
Qy 3361 ATTTTGTCTAGGAGCACTATATGATTTTAAAGTCTATGTTTAAATTAATATGTAATA 3420
Db 3361 ATTTTGTCTAGGAGCACTATATGATTTTAAAGTCTATGTTTAAATTAATATGTAATA 3420
Qy 3421 TTTTTCAGCTATTTTAGTGAATATATTTTATGGTGGGAATAAAAATTTCTACTACAGA 3476
Db 3421 TTTTTCAGCTATTTTAGTGAATATATTTTATGGTGGGAATAAAAATTTCTACTACAGA 3476
```

RESULT 11  
AAT00802

ID AAT00802 standard; DNA; 3475 BP.

XX AAT00802;

XX 19-MAR-1996 (first entry)

XX Human Flk2/flt3 tyrosine kinase receptor gene.

XX Tyrosine kinase receptor; flk2; flt3; agonist; monoclonal antibody;  
KW haematopoiesis; hypoplasemia; anaemia; thrombocytopenia; stem cell; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 58..3039

XX FT /\*tag= a

XX FT /product= flk2/flt3\_tyrosine\_kinase\_receptor

XX W09527062-A1.

XX 12-OCT-1995.

XX 23-MAR-1995; 95WO-US03718.

XX 04-APR-1994; 94US-0222299.

XX (GETH ) GENENTECH INC.

XX Bennett BD, Broz SD, Matthews W, Zeigler FC;

XX WPI, 1995-359636/46.

XX P-PSDB; AAR81869.

XX Agonist antibody which binds to flk2-flt3 tyrosine kinase receptor

- enhances proliferation of haematopoietic stem cells, in the treatment of hypoplasia, anaemia, etc.

Disclosure; Page 41-44; 59pp; English.

DNA (AAT00801) coding for murine stem cell tyrosine kinase receptor flk2 (also called flt3) (AAR1868) was obtained by RT-PCR amplification of RNA isolated from mid-gestation mouse foetal livers using primers based on the murine flt3 sequence, and subcloning of the product into pRK5.1. An flk2/flt3 extracellular domain-IgG1 Fc fusion protein was constructed and used to raise agonist antibodies able to bind to, and activate, flk2/flt3. The human flk2/flt3 receptor gene (AAT00802) and amino acid sequence (AAR81869) are also given.

Sequence 3475 BP; 1042 A; 709 C; 784 G; 940 T; 0 other;

Query Match 99.1%; Score 3470.2; DB 16; Length 3475;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3472; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy	1	CGAGGCGCATCCGAGGGCTGGCGCGGCGCTGGGGACCCGGGCTCCGGAGGCCATG	60
Db	1	CGAGGCGCATCCGAGGGCTGGCGCGGCGCTGGGGACCCGGGCTCCGGAGGCCATG	60
Qy	61	CCGGGCTTGGCGCGGCGGCGGCGGCGGCGGCTGGCTGTTTCTGCAATGATA	120
Db	61	CCGGGCTTGGCGCGGCGGCGGCGGCGGCGGCTGGCTGTTTCTGCAATGATA	120
Qy	121	TTTGGGACTATTCAAAATCAAGATCTGCTGTGATCAAGTGTTTAAATCAATCAAG	180
Db	121	TTTGGGACTATTCAAAATCAAGATCTGCTGTGATCAAGTGTTTAAATCAATCAAG	180
Qy	181	AACATGATTCATCAGTGGGAGTCAATCATATCCATGATCAAGTGTGTTTAAATCAATCAAG	240
Db	181	AACATGATTCATCAGTGGGAGTCAATCATATCCATGATCAAGTGTGTTTAAATCAATCAAG	240
Qy	241	GACCTCGGGTGTGGTTCAGACCCGAGGCTCAGGACAGTGTACGAAGTGCCTGTG	300
Db	241	GACCTCGGGTGTGGTTCAGACCCGAGGCTCAGGACAGTGTACGAAGTGCCTGTG	300
Qy	301	GAATGATGATCTGCTTCCATCACTCACTCACTCACTCACTCACTCACTCACTCACT	360
Db	301	GAATGATGATCTGCTTCCATCACTCACTCACTCACTCACTCACTCACTCACTCACT	360
Qy	361	TCCTGTCTCTGGTCTTTAAGACAGCTCCCTGATTCGCGGCGGCGGCGGCGGCGG	420
Db	361	TCCTGTCTCTGGTCTTTAAGACAGCTCCCTGATTCGCGGCGGCGGCGGCGGCGG	420
Qy	421	AACAGAGGAGTGTCTTCCATGCTCACTTTGAAATGACAGAAACCCAGCTGGAGATAC	480
Db	421	AACAGAGGAGTGTCTTCCATGCTCACTTTGAAATGACAGAAACCCAGCTGGAGATAC	480
Qy	481	CTACTTTTATTCAGAGTGAAGTACCAATATACCAATATTTTACAGTGAATGAAGA	540
Db	481	CTACTTTTATTCAGAGTGAAGTACCAATATACCAATATTTTACAGTGAATGAAGA	540
Qy	541	AATACCTGCTTTTACATTAAGAACCTTACCTTTAGAAAAATGGAACCAAGAGGCC	600
Db	541	AATACCTGCTTTTACATTAAGAACCTTACCTTTAGAAAAATGGAACCAAGAGGCC	600
Qy	601	CTGGTCTGCATATCTGAGAGGCTTCCAGAGCGGATCGTGGATGGGCTTTGGATCA	660
Db	601	CTGGTCTGCATATCTGAGAGGCTTCCAGAGCGGATCGTGGATGGGCTTTGGATCA	660
Qy	661	CAGGGGAAAGCTGTAAAGAGAAAGTCCAGCTGTGTTTAAAGAGGAGGAAAAAGTGCTT	720
Db	661	CAGGGGAAAGCTGTAAAGAGAAAGTCCAGCTGTGTTTAAAGAGGAGGAAAAAGTGCTT	720
Qy	721	CATGAATTTTGGGACGACATTAAGTGCTGTGCTCCAGAAATGAATGGGAGGGAATGC	780
Db	721	CATGAATTTTGGGACGACATTAAGTGCTGTGCTCCAGAAATGAATGGGAGGGAATGC	780
Qy	781	ACCAGGCTGTTTCAATAGATCTAAATCAAACTCCTCAGACCAATTCATTTATTT	840

Db	781	ACCAGGCTGTTTCAATAGATCTAAATCAAACTCCTCAGACCAATTCGCAATTTATTT	840
Qy	841	CTTAAAGTAGGGAAACCCCTTATGGATAAGGTGCAAGCTGTTTCAATGTAACCATGATTC	900
Db	841	CTTAAAGTAGGGAAACCCCTTATGGATAAGGTGCAAGCTGTTTCAATGTAACCATGATTC	900
Qy	901	GGGCTCACTCGGAATTAAGAAACCAAGCACTCGAGGAGGGCAACTTCTTGGATGAGT	960
Db	901	GGGCTCACTCGGAATTAAGAAACCAAGCACTCGAGGAGGGCAACTTCTTGGATGAGT	960
Qy	961	ACCTATTCAAACCAAGCAACTATGATACGGATCTGTTTGTATCATCATGAGTGGCA	1020
Db	961	ACCTATTCAAACCAAGCAACTATGATACGGATCTGTTTGTATCATCATGAGTGGCA	1020
Qy	1021	AGAAAGCAGCGGATATCACTTGTCTTCTTCAAGAGATCCCAAGTCAATCAGCTTTG	1080
Db	1021	AGAAAGCAGCGGATATCACTTGTCTTCTTCAAGAGATCCCAAGTCAATCAGCTTTG	1080
Qy	1081	GTTACCATCGTAGGAAGGGATTTATAATGCTACCAATTCAGTGAAGATTTGAAAT	1140
Db	1081	GTTACCATCGTAGGAAGGGATTTATAATGCTACCAATTCAGTGAAGATTTGAAAT	1140
Qy	1141	GACCAATATGAAGAGTGTGTTTCTGTCAGGTTTAAAGCTTAAAGCTTCAATCAGATCT	1200
Db	1141	GACCAATATGAAGAGTGTGTTTCTGTCAGGTTTAAAGCTTAAAGCTTCAATCAGATCT	1200
Qy	1201	ACGTGGACCTTCTCTCGAAATCAATTTCTTGTGTCAGCAAAAGGGTCTTGATACCGATAC	1260
Db	1201	ACGTGGACCTTCTCTCGAAATCAATTTCTTGTGTCAGCAAAAGGGTCTTGATACCGATAC	1260
Qy	1261	AGCATATCCAAAGTGTGCAATCATAAGCAGCAGGAGGAGATATATATTCATGAGAA	1320
Db	1261	AGCATATCCAAAGTGTGCAATCATAAGCAGCAGGAGGAGATATATATTCATGAGAA	1320
Qy	1321	AATGATGATGCCCAATTTTACCAGGATGTTTACGCTGATATTAAGAGGAACTCAAGTG	1380
Db	1321	AATGATGATGCCCAATTTTACCAGGATGTTTACGCTGATATTAAGAGGAACTCAAGTG	1380
Qy	1381	CTCGCAGAGCATCGGCAAGTCAAGGCTGCTGTTTCTCGGATGGATACCATTCATCT	1440
Db	1381	CTCGCAGAGCATCGGCAAGTCAAGGCTGCTGTTTCTCGGATGGATACCATTCATCT	1440
Qy	1441	TGGAACCTGGAAGAGTGTTCAGACAAAGTCTCCCAATGCAAGAGAGATCACAGAAGA	1500
Db	1441	TGGAACCTGGAAGAGTGTTCAGACAAAGTCTCCCAATGCAAGAGAGATCACAGAAGA	1500
Qy	1501	GTCTGGAATGAAGGCTTACAGAAAGTGTGTCAGTGGGTGTCAGAGGATCTCTA	1560
Db	1501	GTCTGGAATGAAGGCTTACAGAAAGTGTGTCAGTGGGTGTCAGAGGATCTCTA	1560
Qy	1561	AACATGATGAAGCATAAAGGGTCTCTGCTCAAGTGTGTCATCAATTCCTCTGCGC	1620
Db	1561	AACATGATGAAGCATAAAGGGTCTCTGCTCAAGTGTGTCATCAATTCCTCTGCGC	1620
Qy	1621	ACATCTTGTGAGAGTCTCTTTTAAACTCTCCAGGCGCCCTTCCCTTTTCAACAGACAAC	1680
Db	1621	ACATCTTGTGAGAGTCTCTTTTAAACTCTCCAGGCGCCCTTCCCTTTTCAACAGACAAC	1680
Qy	1681	ATCTCATTTATGCAACAAATGGTGTGTTGCTCTCTTCAATGTCGTTTTAAACCTGCTA	1740
Db	1681	ATCTCATTTATGCAACAAATGGTGTGTTGCTCTCTTCAATGTCGTTTTAAACCTGCTA	1740
Qy	1741	ATTTGTCAAGTACAAAGAGCAATTTAGTATGAAGCCAGCTACAGATGGTACAGTG	1800
Db	1741	ATTTGTCAAGTACAAAGAGCAATTTAGTATGAAGCCAGCTACAGATGGTACAGTG	1800
Qy	1801	ACCGGCTCTCAGATATGAGTACTTCTAGCTTGAATTCAGAGAAATGAATATGATCTC	1860
Db	1801	ACCGGCTCTCAGATATGAGTACTTCTAGCTTGAATTCAGAGAAATGAATATGATCTC	1860
Qy	1861	AAATGGAGTTCAGAGAGAAATTTAGAGTTTGGGAGGTACTAGGATCAGGTGCTTTT	1920

1861	DB	AAATGGGAGTTTCCAAGAGAAAAATTTAGATGTTTGGNAAGTACTAGGATCAGTGCCTTTT	1922
1921	QY	GGAAAAAGTGATGAACGCAACGACGTCTATGGAAATTAGCAAAAACAGGAGTCTCAATCCAGGTTT	1980
1921	DB	GGAAAAAGTGATGAACGCAACGACGTCTATGGAAATTAGCAAAAACAGGAGTCTCAATCCAGGTTT	1980
1981	QY	GCCGTCAAATGCTGAAAGAAAAAGCAGACGTCTGAAAGAGAGGCACTCATGTTCAGAA	2040
1981	DB	GCCGTCAAATGCTGAAAGAAAAAGCAGACGTCTGAAAGAGAGGCACTCATGTTCAGAA	2040
2041	QY	CTCAAGATGATGACCCAGCTGGGAAGCCACGAGAAATATCTGAAACTGCTGGGGGGCGTGC	2100
2041	DB	CTCAAGATGATGACCCAGCTGGGAAGCCACGAGAAATATCTGAAACTGCTGGGGGGCGTGC	2100
2101	QY	ACACTGTCAGGACCAATTTACTTTGAATTTTGAATACTGTTGCTATGCTGATCTTCCAAC	2160
2101	DB	ACACTGTCAGGACCAATTTACTTTGAATTTTGAATACTGTTGCTATGCTGATCTTCCAAC	2160
2161	QY	TATCTAAGAGTAAAGAGAAAAATTTACAGAGACTTTGGAACAGAGATTTTCAAGGAAACAC	2220
2161	DB	TATCTAAGAGTAAAGAGAAAAATTTACAGAGACTTTGGAACAGAGATTTTCAAGGAAACAC	2220
2221	QY	AAATTCAGTCTTTTACCCCACTTTTCCAAATCACATCCAAATTTCCAGCAGCTGCTGTCCAAGA	2280
2221	DB	AAATTCAGTCTTTTACCCCACTTTTCCAAATCACATCCAAATTTCCAGCAGCTGCTGTCCAAGA	2280
2281	QY	GAAGTTCAGATACACCCGGACTCGGATCAAACTCTCAGGGCTTCATGGGAATTCATTTCCAC	2340
2281	DB	GAAGTTCAGATACACCCGGACTCGGATCAAACTCTCAGGGCTTCATGGGAATTCATTTCCAC	2340
2341	QY	TCTGAAGATGAAATTTGAATATGAAAAACCAAAAAAGGCTGGAAGAGAGGAGCACTTGAAT	2400
2341	DB	TCTGAAGATGAAATTTGAATATGAAAAACCAAAAAAGGCTGGAAGAGAGGAGCACTTGAAT	2400
2401	QY	GTGCTTACATTTGNAAGATCTTCTTGCTTTGCAATCAAGTTCGCCAAGGAATGGAATTT	2460
2401	DB	GTGCTTACATTTGNAAGATCTTCTTGCTTTGCAATCAAGTTCGCCAAGGAATGGAATTT	2460
2461	QY	CTGGAAATTTAAGTCGTGTGTTACAGAGACCTGCGCCGACAGGAACGTGTGTCAACCCAC	2520
2461	DB	CTGGAAATTTAAGTCGTGTGTTACAGAGACCTGCGCCGACAGGAACGTGTGTCAACCCAC	2520
2521	QY	GGGAAAAGTGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTGATTTCCAAC	2580
2521	DB	GGGAAAAGTGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTGATTTCCAAC	2580
2581	QY	TATGTTGTGAGGGCAATGCCGCTGCTGCTGTAAATTTGGATGGCCCCCGGAAAGCCCTGTTT	2640
2581	DB	TATGTTGTGAGGGCAATGCCGCTGCTGCTGTAAATTTGGATGGCCCCCGGAAAGCCCTGTTT	2640
2641	QY	GAAGGCATCTACACCAATTAAGAGTGATGCTGTGGTCATATGGAATATTACTGTGGGAAATC	2700
2641	DB	GAAGGCATCTACACCAATTAAGAGTGATGCTGTGGTCATATGGAATATTACTGTGGGAAATC	2700
2701	QY	TTCTCACTTGGTGTGAATCTCTTACCCTGGCAATCCGGTTGATGCTAACTCTTCAAAACTG	2760
2701	DB	TTCTCACTTGGTGTGAATCTCTTACCCTGGCAATCCGGTTGATGCTAACTCTTCAAAACTG	2760
2761	QY	ATTCAAAAATGGAAATTTAAATTTGGATCAGGCATTTTATGCTACAGAGAAATATACATTATA	2820
2761	DB	ATTCAAAAATGGAAATTTAAATTTGGATCAGGCATTTTATGCTACAGAGAAATATACATTATA	2820
2821	QY	ATGCAATCTCTCGGCTTTTGAATCTCAAGGAAACGGCCATCTTCCCTAAATTTGACTTCG	2880
2821	DB	ATGCAATCTCTCGGCTTTTGAATCTCAAGGAAACGGCCATCTTCCCTAAATTTGACTTCG	2880
2881	QY	TTTTTTAGGATGTCAAGTGGCAGATGACAGAAAGCGATGTATCAGAAATGTGGATGGCCGT	2940
2881	DB	TTTTTTAGGATGTCAAGTGGCAGATGACAGAAAGCGATGTATCAGAAATGTGGATGGCCGT	2940
2941	QY	GTTTTCGGAATGCTCTCAACCTACCAAAACAGGCGACTTTTCAGCAGAGAGATGGATTTG	3000
2941	DB	GTTTTCGGAATGCTCTCAACCTACCAAAACAGGCGACTTTTCAGCAGAGAGATGGATTTG	3000

Qy	3001	GGGCTACTCTCCGAGGCTCAGGTCGAAGATTTCGTAGAGGAACAATTTAGTTTTAAAGG	3061
Db	3001	GGGCTACTCTCTCCGAGGCTCAGGTCGAAGATTTCGTAGAGGAACAATTTAGTTTTAAAGG	3060
Qy	3061	ACTTCACTCCCTCCACCTATCCCTTAAACAGGCTGTAGATTACCAAAACAAGATTAAATTTTCAT	3120
Db	3061	ACTTCACTCCCTCCACCTATCCCTTAAACAGGCTGTAGATTACCAAAACAAGATTAAATTTTCAT	3120
Qy	3121	CACATAAAGAAAATCTATTATCAACTGCTGCTTCCACAGACTTTTCTTAGAAGCCGTCT	3180
Db	3121	CACATAAAGAAAATCTATTATCAACTGCTGCTTCCACAGACTTTTCTTAGAAGCCGTCT	3180
Qy	3181	GGGTTTACTCTGTTTTCAAGGGACTTTTGTAAAAATCAAATCATCTGTCACAAGGCAG	3240
Db	3181	GGGTTTACTCTGTTTTCAAGGGACTTTTGTAAAAATCAAATCATCTGTCACAAGGCAG	3240
Qy	3241	GAGGAGCTGATAATGAACCTTTATTGGAGCAATTCATCTGCATCCAAGGCCCTTCTCAGGCCG	3300
Db	3241	GAGGAGCTGATAATGAACCTTTATTGGAGCAATTCATCTGCATCCAAGGCCCTTCTCAGGCCG	3300
Qy	3301	GCTTGAGTGAATTTGTGTACCTGAAGTACAGTATATCTTGTAAATACATAAAACAAAAGC	3360
Db	3301	GCTTGAGTGAATTTGTGTACCTGAAGTACAGTATATCTTGTAAATACATAAAACAAAAGC	3360
Qy	3361	ATTTTGTCAAGGAGAGCTAATATGATTTTTTAAAGTCTATGTTTTAAATAATATGTAAA	3420
Db	3361	ATTTTGTCAAGGAGAGCTAATATGATTTTTTAAAGTCTATGTTTTAAATAATATGTAAA	3420
Qy	3421	TTTTTTCAGCTATTAGTGATATATTTTATGGGTGGGAATAAAATTTCTACTACAG	3475
Db	3421	TTTTTTCAGCTATTAGTGATATATTTTATGGGTGGGAATAAAATTTCTACTACAG	3475
RESULT 12			
AAD42484			
ID	AAD42484 standard; cDNA; 3489 BP.		
XX	AAD42484;		
AC	AAD42484;		
XX	15-NOV-2002 (first entry)		
DT	Human receptor protein tyrosine kinase, flk-2 cDNA.		
XX	Human receptor protein tyrosine kinase, flk-2 cDNA.		
DE	Human; receptor protein tyrosine kinase; pTK; haematopoietic cell; growth factor; therapeutic; macrocytic anaemia; aplastic anaemia; bone marrow damage; cancer; chemotherapy; radiation; osteopathic; flk-2; gene; ss.		
KW	Human; receptor protein tyrosine kinase; pTK; haematopoietic cell; growth factor; therapeutic; macrocytic anaemia; aplastic anaemia; bone marrow damage; cancer; chemotherapy; radiation; osteopathic; flk-2; gene; ss.		
KW	Human; receptor protein tyrosine kinase; pTK; haematopoietic cell; growth factor; therapeutic; macrocytic anaemia; aplastic anaemia; bone marrow damage; cancer; chemotherapy; radiation; osteopathic; flk-2; gene; ss.		
KW	Human; receptor protein tyrosine kinase; pTK; haematopoietic cell; growth factor; therapeutic; macrocytic anaemia; aplastic anaemia; bone marrow damage; cancer; chemotherapy; radiation; osteopathic; flk-2; gene; ss.		
XX	Homo sapiens.		
OS	Homo sapiens.		
PH	Key		
FT	CDS		
FT	Location/Qualifiers		
FT	58..3039		
FT	/*tag= a		
FT	/product= "Human flk-2 protein"		
FT	58..138		
FT	/*tag= b		
FT	139..3036		
FT	mat_peptide		
FT	/*tag= c		
FT	/product= "Human mature flk-2 protein"		
FT	139..1689		
FT	/*tag= d		
FT	/note= "Encodes extracellular domain"		
FT	1690..1746		
FT	/*tag= e		
FT	/note= "Encodes transmembrane domain"		
FT	1747..3036		
FT	/*tag= f		
FT	/note= "Encodes intracellular domain"		
XX	US2002072077-A1.		
PN	13-JUN-2002.		
XX	13-JUN-2002.		

XX PF 31-JUL-2001; 2001US-0919408.  
XX PR 19-NOV-1992; 92US-0977451.  
XX PR 30-APR-1993; 93US-0055269.  
XX PR 31-OCT-1994; 94US-0252498.  
XX PR 15-FEB-1996; 96US-0601891.  
XX PR 10-FEB-1998; 98US-0021324.  
XX PR 10-DEC-1998; 98US-0208786.  
XX PR 02-APR-1991; 91US-0679666.  
XX PR 28-JUN-1991; 91US-0728913.  
XX PR 15-NOV-1991; 91US-0793065.  
XX PR 24-DEC-1991; 91US-0813593.  
XX (LEMI/) LEMISCHKA I R.  
XX PA Lemischka IR;  
XX PI WPI; 2002-607237/65.  
XX DR P-PSDB; AAE25819.  
XX DR  
XX PT New protein tyrosine kinase expressed in primitive hematopoietic cells  
XX PT (HC) and not expressed in mature HC, and ligands for the protein, for  
XX PT stimulating proliferation of primitive hematopoietic stem cells  
XX PS Claim 9; Page 22-25; 64pp; English.  
XX CC The present invention relates to receptor protein tyrosine kinase (pTK)  
XX CC expressed in primitive haematopoietic cells and not expressed in mature  
XX CC haematopoietic cells, polynucleotides encoding such proteins and ligands  
XX CC for the protein. Ligands which bind to pTK of the invention are useful  
XX CC for stimulating the proliferation and/or differentiation of primitive  
XX CC mammalian haematopoietic stem cells. The receptor pTK molecules are  
XX CC useful for stimulating the self renewal of the totipotent haematopoietic  
XX CC stem cell and to stimulate the development of all cells of haematopoietic  
XX CC system both in vitro and in vivo. The ligands for the receptors act as  
XX CC haematopoietic growth factors. The ability of the ligands to stimulate  
XX CC proliferation of stem cells both in vitro and in vivo has important  
XX CC therapeutic applications such as treating humans whose primitive stem  
XX CC cells do not sufficiently undergo self-renewal. It is also useful in  
XX CC conditions that occur when defects in haematopoietic stem cells or their  
XX CC related growth factors depress the number of white blood cells such as  
XX CC macrocytic and aplastic anaemia or bone marrow damage resulting from  
XX CC cancer chemotherapy and radiation. The present sequence is a cDNA  
XX CC encoding human receptor pTK, fik-2.  
XX SQ Sequence 3489 BP; 1066 A; 707 C; 780 G; 936 T; 0 other;

Query Match 97.6%; Score 3417; DB 24; Length 3489;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 3489; Conservative 0; Mismatches 0; Indels 12; Gaps 6;

Qy 1 CGAGCGGCATCCGAGGGCTGGGCGGCCCTCGGGGACCCCGGGCTCCGGAGGCCATG 60  
Db 1 CGAGCGGCATCCGAGGGCTGGGCGGCCCTCGGGGACCCCGGGCTCCGGAGGCCATG 60  
Qy 61 CCGGCGTTGGCGCGGACGCGGGGACCGGCTGCTGCTGTTTCTGCAATGATA 120  
Db 61 CCGGCGTTGGCGCGGACGCGGGGACCGGCTGCTGCTGTTTCTGCAATGATA 120  
Qy 121 TTTCGGACTATACAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATGAAG 180  
Db 121 TTTCGGACTATACAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATGAAG 180  
Qy 181 AACAAATGATTCATAGTGGGGAAGTCATCATCATATCCATGGTATCAGATCCCGGAA 240  
Db 181 AACAAATGATTCATAGTGGGGAAGTCATCATCATATCCATGGTATCAGATCCCGGAA 240  
Qy 241 GACCTCGGGTGTGCTTCCAGACCCGAGAGTACAGGACAGTGTACGAAGCTGCCGCTGTG 300  
Db 241 GACCTCGGGTGTGCTTCCAGACCCGAGAGTACAGGACAGTGTACGAAGCTGCCGCTGTG 300  
Qy 301 GAAGTGGATGTATCTGCTTCCATCACTGCAAGTGTGCTGATGCCCGGGAACATT 360

Db 301 GAAGTGGATGTATCTGCTTCCATCACTGCAAGTGTGCTGATGCCCGGGAACATT 360  
Qy 361 TCCTGTCTCTGGGTCTTTAAAGCAGCTCCCTGAAATTCGACGACATTTTGAATTTACAA 420  
Db 361 TCCTGTCTCTGGGTCTTTAAAGCAGCTCCCTGAAATTCGACGACATTTTGAATTTACAA 420  
Qy 421 AACAGAGGAGTGTGTTTCCATGTCTATTTGAAAATCAGAGAAACCCAGCTGAGAAATAC 480  
Db 421 AACAGAGGAGTGTGTTTCCATGTCTATTTGAAAATCAGAGAAACCCAGCTGAGAAATAC 480  
Qy 481 CTACTTTTTTATTCAGAGTCAAGCTCAATATACCAATATTTGTTTACAGTGTGATTAAGA 540  
Db 481 CTACTTTTTTATTCAGAGTCAAGCTCAATATACCAATATTTGTTTACAGTGTGATTAAGA 540  
Qy 541 AATACCTCTCTTTACATTAAGAAGCCTTACTTTTAGAAAAATGAAAAACCCAGGACGCC 600  
Db 541 AATACCTCTCTTTACATTAAGAAGCCTTACTTTTAGAAAAATGAAAAACCCAGGACGCC 600  
Qy 601 CTGGTCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGAATGGGTGCTTTCGGAATCA 660  
Db 601 CTGGTCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGAATGGGTGCTTTCGGAATCA 660  
Qy 661 CAGGGGGAAGCTGTAAAGAGAAAGTCCAGCTGTTGTTTAAAAAGAGGAAAGTGCCTT 720  
Db 661 CAGGGGGAAGCTGTAAAGAGAAAGTCCAGCTGTTGTTTAAAAAGAGGAAAGTGCCTT 720  
Qy 721 CATGAATTTATTTGGGACCGGACATAAAGGTGCTGTGCAGAAAAATGAACCTGGGACGAATGC 780  
Db 721 CATGAATTTATTTGGGACCGGACATAAAGGTGCTGTGCAGAAAAATGAACCTGGGACGAATGC 780  
Qy 781 ACCAGGCTGTTTCAAAATAGATCTAAATCAAACTCTCTCAGACACATTTGCCACAATTTATTT 840  
Db 781 ACCAGGCTGTTTCAAAATAGATCTAAATCAAACTCTCTCAGACACATTTGCCACAATTTATTT 840  
Qy 841 CTTAAAGTAGGGGACCCCTTATGGATAGGTGCAAGCTGTTTCATGTGAACCATGATTC 900  
Db 841 CTTAAAGTAGGGGACCCCTTATGGATAGGTGCAAGCTGTTTCATGTGAACCATGATTC 900  
Qy 901 GGGCTCACCTGGGAATTAGAAAAAAGCAGCTCGAGGAGGCAACTTCTTTTCAGATGAGT 960  
Db 901 GGGCTCACCTGGGAATTAGAAAAAAGCAGCTCGAGGAGGCAACTTCTTTTCAGATGAGT 960  
Qy 961 ACCTATTCAACAAACAGAACTATGATACGGATTCCTGTTTGTATCATCATGCGCA 1020  
Db 961 ACCTATTCAACAAACAGAACTATGATACGGATTCCTGTTTGTATCATCATGCGCA 1020  
Qy 1021 AGAAACGACCGGATACATCTGTTCTCTTCAAAGCATCCAGTCAATCAGCTTTG 1080  
Db 1021 AGAAACGACCGGATACATCTGTTCTCTTCAAAGCATCCAGTCAATCAGCTTTG 1080  
Qy 1081 GTTACCATCGTAGGAAAGGATTTTAAATGCTCAATTCAGTGAAGATTTGAAATTT 1140  
Db 1081 GTTACCATCGTAGGAAAGGATTTTAAATGCTCAATTCAGTGAAGATTTGAAATTT 1140  
Qy 1141 GACCAATATGAAGAGTGTGTTTCTGTCAGGTTTAAAGCCTACCCACAAATCAGATCT 1200  
Db 1141 GACCAATATGAAGAGTGTGTTTCTGTCAGGTTTAAAGCCTACCCACAAATCAGATCT 1200  
Qy 1201 ACGTGGACCTTCTCTCGAAAAATCAATTTCTTGTGAGCAAAAAGGGTCTTTGATAACGGATAC 1260  
Db 1201 ACGTGGACCTTCTCTCGAAAAATCAATTTCTTGTGAGCAAAAAGGGTCTTTGATAACGGATAC 1260  
Qy 1261 AGCATATCCAAAGTGTGCAATCAATGACAGCCAGGAGAGATATATATTTCCATCCAGAA 1320  
Db 1261 AGCATATCCAAAGTGTGCAATCAATGACAGCCAGGAGAGATATATATTTCCATCCAGAA 1320  
Qy 1321 AATGATGATGCCCAATTTTACAAAAATGTTTACGCTGAAATAAAGAAAGAAACCTCAAGTG 1380  
Db 1321 AATGATGATGCCCAATTTTACAAAAATGTTTACGCTGAAATAAAGAAAGAAACCTCAAGTG 1380  
Qy 1381 CTCGAGAAAGCATCGGCAAGTCAAGCTGCTGTTTCTCGGATGGATACCCATTCATCTCT 1440



Db 1381 CTCGAGAAGCATCGGCAAGTCAGCGGCTCTGTTTCTCGGATGGATACCCATTACATCT 1440  
QY 1441 TGGACCTGGAAGAAAGTGTTCAGACAAGTCTCCAACTCGACAAGAGAGATCAAGAAGGA 1500  
Db 1441 TGGACCTGGAGAAGTGTTCAGACAAGTCTCCAACTCGACAAGAGAGATCAAGAAGGA 1500  
QY 1501 GTCTGGAATAGAAAGCGCTAAACAGAAAAGTGTGTTGGACAGTGGGTGTCGAGCAGTACTCTA 1560  
Db 1501 GTCTGGAATAGAAAGCGCTAAACAGAAAAGTGTGTTGGACAGTGGGTGTCGAGCAGTACTCTA 1560  
QY 1561 AACATGCTGAGCAATTAAGAGGTTCTGGTCAAGTGTGTCATACAAATTCCTTGGC 1620  
Db 1561 AACATGCTGAGCAATTAAGAGGTTCTGGTCAAGTGTGTCATACAAATTCCTTGGC 1620  
QY 1621 ACATCTGTGAGACGATCCCTTTAAACTCTCAGGCCCTTCCCTTTCAATCAAGACAAC 1680  
Db 1621 ACATCTGTGAGACGATCCCTTTAAACTCTCAGGCCCTTCCCTTTCAATCAAGACAAC 1680  
QY 1681 ATCTCATTTCTATGCAACAAATGGGTGTGTTGTCCTCTCATTTGTCGTTTAAACCTGCTA 1740  
Db 1681 ATCTCATTTCTATGCAACAAATGGGTGTGTTGTCCTCTCATTTGTCGTTTAAACCTGCTA 1740  
QY 1741 ATTTGTCAACAGTACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGGTG 1800  
Db 1741 ATTTGTCAACAGTACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGGTG 1800  
QY 1801 ACCGGCTCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGAATATGAATATGATCTC 1860  
Db 1801 ACCGGCTCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGAATATGAATATGATCTC 1860  
QY 1861 AAATGGGATTTCCAAAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920  
Db 1861 AAATGGGATTTCCAAAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920  
QY 1921 GGAAAAGTGAAGCAACAGCTTATGGAATTAGCAAAAACAGGAGTCTCAATCAGGTT 1980  
Db 1921 GGAAAAGTGAAGCAACAGCTTATGGAATTAGCAAAAACAGGAGTCTCAATCAGGTT 1980  
QY 1981 GCCGTCAAAA TCGTGAAGAAAAGCAGACAGCTCTGAAAGAGAGCACTCATGTGAGAA 2040  
Db 1981 GCCGTCAAAA TCGTGAAGAAAAGCAGACAGCTCTGAAAGAGAGCACTCATGTGAGAA 2040  
QY 2041 CTCAAGATGATGACCCAGCTGGGAAGCCACGAGAAATTTGTGAACCTGTGGGGCGTGC 2100  
Db 2041 CTCAAGATGATGACCCAGCTGGGAAGCCACGAGAAATTTGTGAACCTGTGGGGCGTGC 2100  
QY 2101 ACACGTGACGACCAATTTACTGTTTGTGATTTGTAATCTGTTGATGTTGTTCTCAAC 2160  
Db 2101 ACACGTGACGACCAATTTACTGTTTGTGATTTTGAATCTGTTGATGTTGTTCTCAAC 2160  
QY 2161 TATCTAAGAGTAAAGAGAAAATTTACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
Db 2161 TATCTAAGAGTAAAGAGAAAATTTACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
QY 2221 AATTTTCAGTTTATACCCCACTTTCCAAATCACATCCAAATTCAGCATGCTGTTCAAGA 2280  
Db 2221 AATTTTCAGTTTATACCCCACTTTCCAAATCACATCCAAATTCAGCATGCTGTTCAAGA 2280  
QY 2281 GAAGTTTCAGATACACCCGAGCTCGGATCAAAATCTCAAGGCTTCAAGGAAATTCATTTCAC 2340  
Db 2281 GAAGTTTCAGATACACCCGAGCTCGGATCAAAATCTCAAGGCTTCAAGGAAATTCATTTCAC 2340  
QY 2341 TCTGAGATGAATTTGAATATGAATATGAATATGAATATGAATATGAATATGAATATGAAT 2400  
Db 2341 TCTGAGATGAATTTGAATATGAATATGAATATGAATATGAATATGAATATGAATATGAAT 2400  
QY 2401 GTGCTTACATTTGAAGATCTTCTTCTGTTGATATCAAGTTTGCCTTCAAGGAAATTT 2460  
Db 2401 GTGCTTACATTTGAAGATCTTCTTCTGTTGATATCAAGTTTGCCTTCAAGGAAATTT 2460  
QY 2461 CTGGAATTTAAGTCGTGTTTTCAGAGACCTGGCCGACAGGAACTGTTGTACCCAC 2520  
Db 2461 CTGGAATTTAAGTCGTGTTTTCAGAGACCTGGCCGACAGGAACTGTTGTACCCAC 2520

RESULT 13  
AAQ49756  
ID AAQ49756 standard; DNA; 3120 BP.

QY 2521 GGGAAAGTGGTGAAGATATGTGAATTTTGGATTTGGCTCGAGATATCATGAGTGAATTCCTCAAC 2580  
Db 2521 GGGAAAGTGGTGAAGATATGTGAATTTTGGATTTGGCTCGAGATATCATGAGTGAATTCCTCAAC 2580  
QY 2581 TATGTTGTTCAGGGCAATGCGGCTGCTCTGCTGTAAATTCGATGCGCCCGGAAAGCCGTGTTT 2640  
Db 2581 TATGTTGTTCAGGGCAATGCGGCTGCTCTGCTGTAAATTCGATGCGCCCGGAAAGCCGTGTTT 2640  
QY 2641 GAAGGCATCTACACCATTAAGAGTGAATCTGCTGCTCATATGGAATATTTACTGTGGGAAATC 2700  
Db 2641 GAAGGCATCTACACCATTAAGAGTGAATCTGCTGCTCATATGGAATATTTACTGTGGGAAATC 2700  
QY 2701 TTCTCATTGTGTGAATCTCTTAACCTGGCAATTCGGTGTGATGCTTCACTTCAAACTG 2760  
Db 2701 TTCTCATTGTGTGAATCTCTTAACCTGGCAATTCGGTGTGATGCTTCACTTCAAACTG 2760  
QY 2761 ATTCAAAATGGAATTTAAATTCGATCAGCCATTTTATGCTACAGAGAAATATACATTTA 2820  
Db 2761 ATTCAAAATGGAATTTAAATTCGATCAGCCATTTTATGCTACAGAGAAATATACATTTA 2820  
QY 2821 ATGCAATCCTGCTGGGCTTTTGAATCAAGGAAACGGCCATCCTTCCCTAAATTTGACTTCG 2880  
Db 2821 ATGCAATCCTGCTGGGCTTTTGAATCAAGGAAACGGCCATCCTTCCCTAAATTTGACTTCG 2880  
QY 2881 TTTTTPAGGATGTCAGCTGGCAGATCAGAGAAGCGATGTATCAGAAATGTGGATGGCCGT 2940  
Db 2881 TTTTTPAGGATGTCAGCTGGCAGATCAGAGAAGCGATGTATCAGAAATGTGGATGGCCGT 2940  
QY 2941 GTTTCGGAATGCTCACAACCTACCAAAACAGGCCACCTTTCAGCAGAGAGATGGAATTTG 3000  
Db 2941 GTTTCGGAATGCTCACAACCTACCAAAACAGGCCACCTTTCAGCAGAGAGATGGAATTTG 3000  
QY 3001 GGGCTACTCTCTCCGACGCTCAGGTGCAAGATTCGTAGAGGAACAATTTAGTTTTPAAGG 3060  
Db 3001 GGGCTACTCTCTCCGACGCTCAGGTGCAAGATTCGTAGAGGAACAATTTAGTTTTPAAGG 3060  
QY 3061 ACTTCATCCCTCCACTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTTTAT 3120  
Db 3061 ACTTCATCCCTCCACTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTTTAT 3120  
QY 3121 CACTAAAAGAAAATCTATATACAACTGCTTCAACGACTTTTCTCTAGAGCGGCTCT 3180  
Db 3121 CACTAAAAGAAAATCTATATCAA--GCTGCTTCAACGACTTTTCTCTAGAGCGGCTCT 3180  
QY 3181 GGGTTTACTCTTGTGTTTCAAAAGGAGCTTTTGTAAATCAAAATCAATCCTCTCAAGGCGAG 3240  
Db 3181 GGGTTTACTCTTGTGTTTCAAAAGGAGCTTTTGTAAATCAAAATCAATCCTCTCAAGGCGAG 3240  
QY 3179 GGGTTTACTCTTGTGTTTCAAAAGG--TTTGTAAATCAAAATCAATCCTCTCAAGGCGAG 3236  
Db 3179 GGGTTTACTCTTGTGTTTCAAAAGG--TTTGTAAATCAAAATCAATCCTCTCAAGGCGAG 3236  
QY 3241 GAGGAGCTGATAATGAACTTTTATGAGGCAATTTGATCTGATCCAAAGGCTTCTCAGGCGG 3300  
Db 3241 GAGGAGCTGATAATGAACTTTTATGAGGCAATTTGATCTGATCCAAAGGCTTCTCAGGCGG 3300  
QY 3301 GCTTCAGTGAATGTTGTACCTGAAGTACAGTATATCTTGTAAATACATAAAACAAGG 3360  
Db 3301 GCTTCAGTGAATGTTGTACCTGAAGTACAGTATATCTTGTAAATACATAAAACAAGG 3360  
QY 3352 GCTTCAGTGAATGTTGTACCTGAAGTACAGTATATCTTGTAAATACATAAAACAAGG 3352  
Db 3352 GCTTCAGTGAATGTTGTACCTGAAGTACAGTATATCTTGTAAATACATAAAACAAGG 3352  
QY 3361 AATTTGCTTAAAGAGAGCTTAATATGATTTTAACTCTATGTTTAAATATATATGTA 3420  
Db 3361 AATTTGCTTAAAGAGAGCTTAATATGATTTTAACTCTATGTTTAAATATATATGTA 3420  
QY 3353 AATTTGCTTAAAGAGAGCTTAATAT--TTTAAAGTCTATGTTTAAATATATATGTA 3410  
Db 3353 AATTTGCTTAAAGAGAGCTTAATAT--TTTAAAGTCTATGTTTAAATATATATGTA 3410  
QY 3421 TTTTTCAGCTATTTTATGATATATTTTATGCTGGGATTAATTTTCTACTACAGAAAAA 3480  
Db 3421 TTTTTCAGCTATTTTATGATATATTTTATGCTGGGATTAATTTTCTACTACAGAAAAA 3480  
QY 3481 AAAAAA 3501  
Db 3489 AAAAAA 3489



XX AAQ49756;  
AC XX  
DT 25-MAR-2003 (updated)  
DT 10-MAR-1994 (first entry)  
XX  
DE PTK gene LpTK25.  
XX  
XX PTK; protein tyrosine kinase; catalytic domain; c-kit; megakaryocyte;  
KW lymphocyte; amplification; primer; polymerase chain reaction; PCR; ds.  
XX  
XX Homo sapiens.  
XX  
XX WO9315201-A1.  
XX  
XX 05-AUG-1993.  
XX  
XX 22-JAN-1993; 93WO-US00586.  
XX  
XX 22-JAN-1992; 92US-0826935.  
XX  
XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.  
XX  
XX Avraham H, Cowley S, Groopman J, Scadden D;  
XX  
XX WPI; 1993-320330/40.  
XX  
XX New protein tyrosine kinase genes and proteins encoded by genes -  
XX are of human mega-karyocytic origin  
XX  
XX Claim 2; Fig 7; 60pp; English.  
XX  
XX PTK genes were identified using two sets of degenerative  
CC oligonucleotide primers: a first set which amplifies all PTK DNA  
CC segments (AAQ49743-44), and a second set which amplifies highly  
CC conserved sequences present in the catalytic domain of the c-kit  
CC subgroup of PTKs (AAQ49745-46). The PTK genes identified are described  
CC in AAQ49747-57 and AAR41897-02.  
CC The LpTKs are expressed in lymphocytic cells, as well as  
CC megakaryocytic cells.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX  
XX Sequence 3120 BP; 945 A; 643 C; 699 G; 833 T; 0 other;  
SQ

Query Match 88.2%; Score 3086.4; DB 14; Length 3120;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 3099; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 58 ATGCGCGGTTGGCGCGGACGCGGCGACCGCGCTGCTGCTGTTTCTTCTGCAATG 117  
DB 1 ATGAGAGCGTTGGCGCGGACGCGGCGGCGGCGGCGGCTGCTGCTGTTTCTGCAATG 60

QY 118 ATATTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATCAT 177  
DB 61 ATATTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATCAT 120

QY 178 AAGAACAAATCATCATAGTGGGAAAGTCAATCATATCCCATGGTATCAGAAATCCCG 237  
DB 121 AAGAACAAATCATCATAGTGGGAAAGTCAATCATATCCCATGGTATCAGAAATCCCG 180

QY 238 GAAGACCTCCGGGTGTGGTTGGAGACCCAGAGCTCAGGACAGGTACGAAGCTGCCGCT 297  
DB 181 GAAGACCTCCGGGTGTGGTTGGAGACCCAGAGCTCAGGACAGGTACGAAGCTGCCGCT 240

QY 298 GTGGAAGTGGATGATCTGCTTCCATCACACTGCAAGTGTGTCATGCCAGGGAAC 357  
DB 241 GTGGAAGTGGATGATCTGCTTCCATCACACTGCAAGTGTGTCATGCCAGGGAAC 300

QY 358 ATTTCTGCTCTGGGCTTTAAGCAGAGCTCCCTGAAATGGCCAGCACAATTTGATTTA 417  
DB 301 ATTTCTGCTCTGGGCTTTAAGCAGAGCTCCCTGAAATGGCCAGCACAATTTGATTTA 360

QY 418 CAANAACAGAGAGGTGTTTCCATGTTTTCGAAATGACAAACCCCAAGCTGGAGAA 477

DB 361 CAANAACAGAGAGGTGTTTCCATGTTTTCGAAATGACAAACCCCAAGCTGGAGAA 420

QY 478 TACCTACTTTTATTACAGAGTGAAGCTACCAATTAACAATATTGTTTACAGTGAATATA 537

DB 421 TACCTACTTTTATTACAGAGTGAAGCTACCAATTAACAATATTGTTTACAGTGAATATA 480

QY 538 AGAATATCCCTGCTTTTACATTAAGAAGACCTTACTTTAGAAAATGGAACCCAGGAC 597

DB 481 AGAATATCCCTGCTTTTACATTAAGAAGACCTTACTTTAGAAAATGGAACCCAGGAC 540

QY 598 GCCCTGGTCTGATATCTGAGAGCGTTCCAGAGCGATCGTGAAGTGGTCTTTCGGAT 657

DB 541 GCCCTGGTCTGATATCTGAGAGCGTTCCAGAGCGATCGTGAAGTGGTCTTTCGGAT 600

QY 658 TCACAGGGGAAAGCTGTAAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAAAGTG 717

DB 601 TCACAGGGGAAAGCTGTAAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAAAGTG 660

QY 718 CTTTCAATGAATTTTGGGACGACATAAGTCTGTGCGAAGAAATGAATCGGCGAGGAA 777

DB 661 CTTTCAATGAATTTTGGGACGACATAAGTCTGTGCGAAGAAATGAATCGGCGAGGAA 720

QY 778 TGCACAGGCTGTTTCACAATAGATCTAAATCAAACTCCTCAGACCAATTCGCAATTA 837

DB 721 TGCACAGGCTGTTTCACAATAGATCTAAATCAAACTCCTCAGACCAATTCGCAATTA 780

QY 838 TTTCTTAAAGTAGGGGAAACCTTTATGATAAGGTGCAAGCTGTTTCAATGTAACCAATGA 897

DB 781 TTTCTTAAAGTAGGGGAAACCTTTATGATAAGGTGCAAGCTGTTTCAATGTAACCAATGA 840

QY 898 TTCGGGCTCCTCTGGGAATTAAGAAACAAAGCACTCGAGAGGGCAACTTCTTTGAGATG 957

DB 841 TTCGGGCTCCTCTGGGAATTAAGAAACAAAGCACTCGAGAGGGCAACTTCTTTGAGATG 900

QY 958 AGTACCTATTCAACAAACAGAACTATGATCGGATCTGTTGCTTTTGTATCATCAGTG 1017

DB 901 AGTACCTATTCAACAAACAGAACTATGATCGGATCTGTTGCTTTTGTATCATCAGTG 960

QY 1018 GCAAGAAACGACACCGGATCTACACTTGTGTTCTTCAAGAGTCCCAAGTCAATCAGCT 1077

DB 961 GCAAGAAACGACACCGGATCTACACTTGTGTTCTTCAAGAGTCCCAAGTCAATCAGCT 1020

QY 1078 TTGTTTACCATCTGAGAAAGGATTTATAAATGCTACCAATTCAGTGAAGATTAATGAA 1137

DB 1021 TTGTTTACCATCTGAGAAAGGATTTATAAATGCTACCAATTCAGTGAAGATTAATGAA 1080

QY 1138 ATTGACCAATATGAAGATTTTGTGTTTCTGTCAGGTTTAAAGCTACCCCAATCAGA 1197

DB 1081 ATTGACCAATATGAAGATTTTGTGTTTCTGTCAGGTTTAAAGCTACCCCAATCAGA 1140

QY 1198 TGTACGTGGACCTTCTCTCGAAATCATTTCTTGTGAGCAAAAGGTTCTTGATACCGA 1257

DB 1141 TGTACGTGGACCTTCTCTCGAAATCATTTCTTGTGAGCAAAAGGTTCTTGATACCGA 1200

QY 1258 TACAGCATATCCAAATTTTGAATTCATAAGCACAGCCAGAGAGATATATATTCATGCA 1317

DB 1201 TACAGCATATCCAAATTTTGAATTCATAAGCACAGCCAGAGAGATATATATTCATGCA 1260

QY 1318 GAAATATGATGATGCCCAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGA 1377

DB 1261 GAAATATGATGATGCCCAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGA 1320

QY 1378 GTCTCTGCAAGACATCGGCAAGTCCAGGCTGCTGTTTCTCGATGATACCCATTTACCA 1437

DB 1321 GTCTCTGCAAGACATCGGCAAGTCCAGGCTGCTGTTTCTCGATGATACCCATTTACCA 1380

QY 1438 TCTTGGACCTGGAAGAGTGTTCAGACCAAGTCTCCCACTGCAAGAGAGATCACAGAA 1497

DB 1381 TCTTGGACCTGGAAGAGTGTTCAGACCAAGTCTCCCACTGCAAGAGAGATCACAGAA 1440

QY 1498 GGAGTCTGGATAGAAAGGCTAAAGAAAGTGTGTTGACAGTGGGTGTCGAGCAGTACT 1557

Db	1441	GGAGTCTGGAATAGAAAGGCTAACAGAAAAGTGTGTTGGACAGTGGGTGCGAGCAGTACT	1500
Qy	1558	CTAAACAATGAGTGAAGCCATAAAGGTTCTCTGGTCAAGTGTCTGTGATACATAAATCCCTTT	1617
Db	1501	CTAAACATGAGTGAAGCCATAAAGGTTCTCTGGTCAAGTGTCTGTGATACATAAATCCCTTT	1560
Qy	1618	GGCACA TCTTGTGAGACGATCTCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAGAC	1677
Db	1561	GGCACATCTGTGAGACGATCTCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAGAC	1620
Qy	1678	AACATCTCAATCTATGCAACAAATGGTGTGTTGTTCTCTCTCTCAATGTCGTTTAAACCCCTG	1737
Db	1621	AACATCTCAATCTATGCAACAAATGGTGTGTTGTTCTCTCTCTCAATGTCGTTTAAACCCCTG	1680
Qy	1738	CTAATTTGTCAACAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAG	1797
Db	1681	CTAATTTGTCAACAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAG	1740
Qy	1798	GTGACCGGCTCTCAGATAA TGAAGTACTTCTACGTTGATTTTCAGAGAAATATGAATATGAT	1857
Db	1741	GTGACCGGATCTCTCAGATAATGATGACTTCTACGTTGATTTTCAGAGAAATATGAATATGAT	1800
Qy	1858	CTCAATTTGGAGTTTCCAGAGAAAATTTAGAGTTTGGAGGTACTAGGATCAGGTGCT	1917
Db	1801	CTCAATTTGGAGTTTCCAGAGAAAATTTAGAGTTTGGAGGTACTAGGATCAGGTGCT	1860
Qy	1918	TTTGGAAAAGTGTGAACGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAG	1977
Db	1861	TTTGGAAAAGTGTGAACGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAG	1920
Qy	1978	GTTTGGCCCTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCA	2037
Db	1921	GTTTGGCCCTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCA	1980
Qy	2038	GAACTCAAGATGATGACCCAGCTGGGAGGCAACAGAGATATTTGTAACCTGCTGGGGCG	2097
Db	1981	GAACTCAAGATGATGACCCAGCTGGGAGGCAACAGAGATATTTGTAACCTGCTGGGGCG	2040
Qy	2098	TGCACATGTCAGACCAATTTACTGATTTTTCGAATCTGCTATGTTGATCTTCTC	2157
Db	2041	TGCACATGTCAGACCAATTTACTGATTTTTCGAATCTGCTATGTTGATCTTCTC	2100
Qy	2158	AACATATCTAAGAAATGAAAGAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAA	2217
Db	2101	AACATATCTAAGAAATGAAAGAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAA	2160
Qy	2218	CACAAATTTCAATTTTACCCCACTTTTCAATCAGATCCTCAATTTCCAGATGCTGTTCA	2277
Db	2161	CACAAATTTCAATTTTACCCCACTTTTCAATCAGATCCTCAATTTCCAGATGCTGTTCA	2220
Qy	2278	AGAGAAGTTTCAGATACACCCGAGCTCGGATCAAAATCTCAGGCTTCA TGGGAAATTCATTT	2337
Db	2221	AGAGAAGTTTCAGATACACCCGAGCTCGGATCAAAATCTCAGGCTTCA TGGGAAATTCATTT	2280
Qy	2338	CACCTCTGAAGATGAAATGAAATGAAACCAAAAAGGCTGGAAGAGAGAGGAGACTTG	2397
Db	2281	CACCTCTGAAGATGAAATGAAATGAAACCAAAAAGGCTGGAAGAGAGAGGAGACTTG	2340
Qy	2398	AATGTGCTTACATTTGAAAGATCTTCTTTGCTTTTGCATATCAAGTTGCCAAAGGATGAA	2457
Db	2341	AATGTGCTTACATTTGAAAGATCTTCTTTGCTTTTGCATATCAAGTTGCCAAAGGATGAA	2400
Qy	2458	TTTCTGGAAATTTAAGTCTGTGTTTTCAGAGAGACTGGCCGCGCAGGAACTGCTGTGACC	2517
Db	2401	TTTCTGGAAATTTAAGTCTGTGTTTTCAGAGAGACTGGCCGCGCAGGAACTGCTGTGACC	2460
Qy	2518	CACGGAAAGTGGTGAAGATATGATCTTTGATTTGGCTTCGAGATATCATAGTATGATCC	2577
Db	2461	CACGGAAAGTGGTGAAGATATGATCTTTGATTTGGCTTCGAGATATCATAGTATGATCC	2520
Qy	2578	AACATATGTTGTGAGGGCAATGCCCGTCTGCTGTTAAATGATGAGCCGCCCAAGGCTG	2637
Db	2521	AACATATGTTGTGAGGGCAATGCCCGTCTGCTGTTAAATGATGAGCCGCCCAAGGCTG	2580
Qy	2638	TTTGAAGGCATCTACACCAATTAAGAGTGTCTGCTCATATGGAATATTACTGTGGGAA	2697
Db	2581	TTTGAAGGCATCTACACCAATTAAGAGTGTCTGCTCATATGGAATATTACTGTGGGAA	2640
Qy	2698	ATCTTCTCACTTGGTGTGAATCTTACCCTGGCATTTCCGGTTGATGCTAACTTCTACAA	2757
Db	2641	ATCTTCTCACTTGGTGTGAATCTTACCCTGGCATTTCCGGTTGATGCTAACTTCTACAA	2700
Qy	2758	CTGATTTCAAAATGGATTTTAAATGGATCAGCCATTTTATGCTACAGAGAAATATACATT	2817
Db	2701	CTGATTTCAAAATGGATTTTAAATGGATCAGCCATTTTATGCTACAGAGAAATATACATT	2760
Qy	2818	ATATGCAATCTGCTGGGCTTTTGAATCAAGGAAAGGCCATCTCTCCCTAATTTGACT	2877
Db	2761	ATATGCAATCTGCTGGGCTTTTGAATCAAGGAAAGGCCATCTCTCCCTAATTTGACT	2820
Qy	2878	TCGTTTTTATAGGATGTCAGCTGGCAGATGCAAGAGAAAGCCATGATCAGAAATGTGGATGGC	2937
Db	2821	TCGTTTTTATAGGATGTCAGCTGGCAGATGCAAGAGAAAGCCATGATCAGAAATGTGGATGGC	2880
Qy	2938	CGTGTTCGGAATGTCTCTCACCTACCAAAAACAGGCGACCTTTTACGACAGAGATGGAT	2997
Db	2881	CGTGTTCGGAATGTCTCTCACCTACCAAAAACAGGCGACCTTTTACGACAGAGATGGAT	2940
Qy	2998	TTGGGGCTACTCTCTCCGCAAGGCTCAGGTGCAAGATTCGTAGAGGAA CAATTTAGTTTAA	3057
Db	2941	TTGGGGCTACTCTCTCCGCAAGGCTCAGGTGCAAGATTCGTAGAGGAA CAATTTAGTTTAA	3000
Qy	3058	AGGACTTCACTCCCTCCACCTATCCCTTAA CAGGCTGTAGATTACCAAAAACAGGATTAATTT	3117
Db	3001	AGGACTTCACTCCCTCCACCTATCCCTTAA CAGGCTGTAGATTACCAAAAACAGGATTAATTT	3060
Qy	3118	CATCACTAAAAGAAAATCTATTATCAACTGCTGCTTCA CCGACTTTTCTCTAGAGGCG	3177
Db	3061	CATCACTAAAAGAAAATCTATTATCAACTGCTGCTTCA CCGACTTTTCTCTAGAGGCG	3120
RESULT 14			
AAT03096			
ID	AAT03096 standard; DNA; 3120 BP.		
XX	AAT03096;		
AC	AAT03096;		
XX	14-FEB-1996 (first entry)		
DT	Protein tyrosine-kinase LpTK25 gene.		
XX	Protein tyrosine-kinase; pTK; LpTK25; agonist; cell growth;		
KW	differentiation; ss.		
XX	Homo sapiens.		
OS	Homo sapiens.		
XX	W09527061-AL.		
FN	12-OCT-1995.		
XX	04-APR-1995; 95WO-US04228.		
PF	04-APR-1994; 94US-0222616.		
XX	(GETH ) GENENTECH INC.		
PA	Bennett BD, Goeddel D, Lee JM, Matthews W, Tsai SP;		
XX	Wood WI;		
PI	WPI; 1995-366160/47.		
XX	Agonist antibodies which activate specific protein tyrosine		
PT	kinase(s) - also activate chimeric proteins of kinase extracellular		
PT	domain and Ig constant domain, useful for studying, and therapeutic		
XX	modulation of, cell growth and differentiation		



Db	1921	GT	TACCGTCAAAATGCTGAAAGAAAAAGACAGACAGCTCTGAAAGAGAGGAGGACCTCATGTCA	1989
Qy	2038	GA	ACTCAAGATGATGACCCAGCTGGGAAGCCACAGAGAAATATGTGAACCTGCTCGGGGCG	2097
Db	1981	GA	CTCAAGATGATGACCCAGCTGGGAAGCCACAGAGAAATATGTGAACCTGCTCGGGGCG	2040
Qy	2098	TG	CACATGTCAGGACCAATTTATCTTGATTTTTTGAATACTGTTGCTATGTTGATCTTCTC	2157
Db	2041	TG	CACATGTCAGGACCAATTTATCTTGATTTTTTGAATACTGTTGCTATGTTGATCTTCTC	2100
Qy	2158	AA	CTATCTAAGAACTAAAGAGAAAAATTTCCACAGGACTTGGACAGAGATTTTCAACGGAA	2217
Db	2101	AA	CTATCTAAGAGTAAAGAGAAAAATTTCCACAGGACTTGGACAGAGATTTTCAAGAA	2160
Qy	2218	CA	CAATTTTCACTTTTATCCCCACCTTTCCAAATCACATCCAAATCCAGCATCCCTGGTTCA	2277
Db	2161	CACA	TTTCACTTTTATCCCCACCTTTCCAAATCACATCCAAATCCAGCATCCCTGGTTCA	2220
Qy	2278	AG	AAGATTTAGATACACCCGGACTCGGATCAAATCTCAGGGCTTCAITGGGAATTTCAATTT	2337
Db	2221	AG	AAGATTTAGATACACCCGGACTCGGATCAAATCTCAGGGCTTCAITGGGAATTTCAATTT	2280
Qy	2338	CAC	CTGAAGATGAATTTGAATATGAATAACCAAAAAGGCTGGAGAGAGGAGGACTTG	2397
Db	2281	CAC	CTGAAGATGAATTTGAATATGAATAACCAAAAAGGCTGGAGAGAGGAGGACTTG	2340
Qy	2398	AA	TGCTTTACATTTGAAGATCTTCTTTGCTTTGCATATCAAGTTGCCAAAGGAATGGAA	2457
Db	2341	AA	TGCTTTACATTTGAAGATCTTCTTTGCTTTGCATATCAAGTTGCCAAAGGAATGGAA	2400
Qy	2458	TTT	CTGGAATTTAAGTCGTGTGTTTACAGAGACTGCGCGCCAGGAACGTGTTGTCAAC	2517
Db	2401	TTT	CTGGAATTTAAGTCGTGTGTTTACAGAGACTGCGCGCCAGGAACGTGTTGTCAAC	2460
Qy	2518	CAC	GGAAGTGGTGAAGATATGTACTTTGGATTTGGCTCGAGATATCATGAGTGATTC	2577
Db	2461	CAC	GGAAGTGGTGAAGATATGTACTTTGGATTTGGCTCGAGATATCATGAGTGATTC	2520
Qy	2578	AA	CTATGTTGTACGGGGCAATGCCCGTCTGCTCTAATAATGGATGGCGCCCGAAAGCCTG	2637
Db	2521	AA	CTATGTTGTACGGGGCAATGCCCGTCTGCTCTAATAATGGATGGCGCCCGAAAGCCTG	2580
Qy	2638	TTT	GAGGCATCTACCATTTAAGATGATGTCTGGTCAATAGGAAATTTATCTGTGGAA	2697
Db	2581	TTT	GAGGCATCTACCATTTAAGATGATGTCTGGTCAATAGGAAATTTATCTGTGGAA	2640
Qy	2698	AT	CTTCTCACTTGGTGTGAATCCTTACCTGGCAATTCGGTTGATGCTAACTTCTACAAA	2757
Db	2641	AT	CTTCTCACTTGGTGTGAATCCTTACCTGGCAATTCGGTTGATGCTAACTTCTACAAA	2700
Qy	2758	CT	GATTCAAAATGATTTTAAAATGGATCAGGCATTTTATGCTACAGAAAGAAATATACATTT	2817
Db	2701	CT	GATTCAAAATGATTTTAAAATGGATCAGGCATTTTATGCTACAGAAAGAAATATACATTT	2760
Qy	2818	ATA	TGCAATCTCTCGTGGCTTTTGAATCAAGGAAACGGCCATCTCTTCCCTTAATTTGACT	2877
Db	2761	ATA	TGCAATCTCTCGTGGCTTTTGAATCAAGGAAACGGCCATCTCTTCCCTTAATTTGACT	2820
Qy	2878	TC	TTTTTATGATGTCACTGGCAGATGTCAGAGAAAGCGATGATCAGAAATGTGGATGGC	2937
Db	2821	TC	TTTTTATGATGTCACTGGCAGATGTCAGAGAAAGCGATGATCAGAAATGTGGATGGC	2880
Qy	2938	CG	TGTTTCGGAATGCTCTCACCTACACCTACCAAAACAGGCGACTTTTCACAGAGATGGAT	2997
Db	2881	CG	TGTTTCGGAATGCTCTCACCTACACCTACCAAAACAGGCGACTTTTCACAGAGATGGAT	2940
Qy	2998	TT	GGGCTACTCTCTCCGCGAGGCTCAGGTCGAAATTCGTAGAGGAAACAATTTAGTTTAA	3057
Db	2941	TT	GGGCTACTCTCTCCGCGAGGCTCAGGTCGAAATTCGTAGAGGAAACAATTTAGTTTAA	3000
Qy	3058	AG	CACTTACCCCTCAACTATCCCTTAAACAGGCTGTAGATTTACAAAACAAGATTAAATTT	3117
Db	3000	AG	CACTTACCCCTCAACTATCCCTTAAACAGGCTGTAGATTTACAAAACAAGATTAAATTT	3060

QY	3118	CATCACTAAAGAAAATCTATTATCAACTCTCTCTCTCAACAGACTTTTCTCTAGAGCGG	3117
Db	3061	CATCACTAAAGAAAATCTATTATCAACTCTCTCTCAACAGACTTTTCTCTAGAGCGG	3120
RESULT 15			
AAV39041	ID AAV39041 standard; cDNA to mRNA; 2949 BP.		
XX	AC	AAV39041;	
XX	DT	12-OCT-1998 (first entry)	
XX	DE	Human receptor type protein kinase FLT3 encoding cDNA SEQ ID NO:24.	
XX	KW	Human; receptor type protein kinase; FLT3; diagnosis; leukaemia;	
XX	KW	membrane-vincinal region; tyrosine kinase; juxtamembrane region; ds.	
XX	OS	Homo sapiens.	
XX	PH	Key	Location/Qualifiers
FT	CDS	1..2949	
FT		/*tag= a	
FT		/note= "no stop codon given"	
XX	PN	W09817808-A1.	
XX	PD	30-APR-1998.	
XX	PF	13-OCT-1997; 97WO-JP03667.	
XX	PR	18-OCT-1996; 96JP-0297329.	
XX	PA	(TAKI ) TAKARA SHUZO CO LTD.	
XX	PI	Yokota S;	
XX	WPI	1998-362333/31.	
DR	P-PSDB	AAW63588.	
XX			
PT		Nucleic acid sequences encoding receptor type protein kinase -	
PT		useful for diagnosing adult acute leukaemia of, e.g. FAB sub-classes	
PT	M2		
XX			
PS	Claim 7;	Page 62-64; 80pp; Japanese.	
XX			
CC		New nucleic acid sequences have been isolated which encode receptor type	
CC		protein kinases (especially a tyrosine kinases) having tandem repeats in	
CC		the juxtamembrane region. Also described in the present invention are:	
CC		(1) any one of 10 amino acid (aa) sequences (AAW63580 to AAW63589)	
CC		representing FLT3 juxtamembrane receptor type protein kinases found in	
CC		leukaemia patients of FAB subclass M2, M4 or M5 (nucleic acid sequences	
CC		encoding the proteins are given in AAV39028 to AAV39042); (2) antibodies	
CC		recognising the kinases or their portions including the tandem repeat	
CC		region; (3) nucleic acid sequences hybridising with the nucleic acids;	
CC		(4) a method for detecting the nucleic acid sequences in human tissue	
CC		samples by: (i) extracting nucleic acid; (ii) amplifying the kinase	
CC		gene, and (iii) determining the size of the gene for comparing the size	
CC		of the normal gene not containing tandem repeats, and (5) kits for	
CC		carrying out the detection. The products and methods may be used for	
CC		diagnosing adult acute leukaemia of FAB subclasses M2, M4 and M5.	
XX			
SQ	Sequence 2949 BP;	895 A; 603 C; 670 G; 781 T; 0 other;	
	Query Match	82.1%;	Score 2875.8; DB 19; Length 2949;
	Best Local Similarity	98.9%;	Pred. No. 0;
	Matches 2917;	Conservative	0; Mismatches 2; Indels 30; Gaps 1;
QY	58	ATGCCGCGTTGCGCGCGACGCGGCGACCGTCCCTGCTCTGTTCTTTCTGCAATG	117
Db	1	ATGCCGCGTTGCGCGCGACGCGGCGACCGTCCCTGCTCTGTTCTTTCTGCAATG	60

118 ATATTTGGGACTATTACAAATCAGATCTGCTGTGATCAAGTGTGTTTTTAATCAATCAT 177  
119 |||||  
120 61 ATATTTGGGACTATTACAAATCAGATCTGCTGTGATCAAGTGTGTTTTTAATCAATCAT 120  
121 |||||  
178 AAGAAATGATGATTCATCAGTGGGAAAGTCATCATATCCATCGTATCAGAAATCCCG 237  
179 |||||  
180 121 AAGAAATGATGATTCATCAGTGGGAAAGTCATCATATCCATCGTATCAGAAATCCCG 180  
181 |||||  
238 GAAGACCTCGGGTGTGCGTTGAGACCCAGAGCTCAGGACAGTGTA CGAAGCTGCCGCT 297  
239 |||||  
181 GAAGACCTCGGGTGTGCGTTGAGACCCAGAGCTCAGGACAGTGTA CGAAGCTGCCGCT 240  
241 |||||  
298 GTGGAAGTGGATGATCTGCTTCCATCACACTGCAAGTGTGCTGATGCCCCAGGGAAC 357  
299 |||||  
241 GTGGAAGTGGATGATCTGCTTCCATCACACTGCAAGTGTGCTGATGCCCCAGGGAAC 300  
301 |||||  
358 ATTTCTCTGTCTCTGGGCTTTTAAAGCACAGCTCCCTGAAATGCCAGCCACATTTTGATTTA 417  
359 |||||  
301 ATTTCTCTGTCTCTGGGCTTTTAAAGCACAGCTCCCTGAAATGCCAGCCACATTTTGATTTA 360  
361 |||||  
418 CAAAAACAGAGAGTGTGTTTCCATGGTCAATTTTGAATAATGACAGAAACCCAGCTGGAGAA 477  
419 |||||  
361 CAAAAACAGAGAGTGTGTTTCCATGGTCAATTTTGAATAATGACAGAAACCCAGCTGGAGAA 420  
421 |||||  
478 TACCTACTTTTATTACAGAGTGAAGCTTACCAATTTTACCAATATGTTTACAGTGAGTATA 537  
479 |||||  
421 TACCTACTTTTATTACAGAGTGAAGCTTACCAATTTTACCAATATGTTTACAGTGAGTATA 480  
481 |||||  
538 AGAAATACCCCTGCTTTTACACATTAAGAGACCTTACTTTAGRAAAATGGAACACAGGAC 597  
539 |||||  
481 AGAAATACCCCTGCTTTTACACATTAAGAGACCTTACTTTAGRAAAATGGAACACAGGAC 540  
541 |||||  
598 GCCTCTGCTGTCATATCTGAGAGCGTTTCCAGAGCCGATCGTGGAAATGGGCTTTGCGAT 657  
599 |||||  
541 GCCTCTGCTGTCATATCTGAGAGCGTTTCCAGAGCCGATCGTGGAAATGGGCTTTGCGAT 600  
601 |||||  
658 TCAACAGGGGAAAGCTGTAAAGAGAAAGTCCAGCTGTTGTTTAAAGAGGAAAGTG 717  
659 |||||  
601 TCAACAGGGGAAAGCTGTAAAGAGAAAGTCCAGCTGTTGTTTAAAGAGGAAAGTG 660  
661 |||||  
718 CTTTATGAATTTATTCGGAGCGACATAAGTGTGTCGACAGAAATGAACTGGGACGGAA 777  
719 |||||  
661 CTTTATGAATTTATTCGGAGCGACATAAGTGTGTCGACAGAAATGAACTGGGACGGAA 720  
721 |||||  
778 TGACACAGGCTGTTCACATAGATCTAAATCAAACCTCTCAGACCCACATTCGCCACATTA 837  
779 |||||  
721 TGACACAGGCTGTTCACATAGATCTAAATCAAACCTCTCAGACCCACATTCGCCACATTA 780  
781 |||||  
838 TTTCTTAAAGTAGGGGAAACCTTTATGGATAAGGTGCAAAAGCTGTTCAATGTGAACCATGGA 897  
839 |||||  
781 TTTCTTAAAGTAGGGGAAACCTTTATGGATAAGGTGCAAAAGCTGTTCAATGTGAACCATGGA 840  
841 |||||  
898 TTTGGGCTCACCTGGGAAATAGAAACAAAGCACTCGAGAGGGCAACTCTTTGAGATG 957  
899 |||||  
841 TTTGGGCTCACCTGGGAAATAGAAACAAAGCACTCGAGAGGGCAACTCTTTGAGATG 900  
901 |||||  
958 AGTACTATTCAACAAACAGAACTATGATACCGATTCTGTTGCTTTTGTATCATCAGTG 1017  
959 |||||  
901 AGTACTATTCAACAAACAGAACTATGATACCGATTCTGTTGCTTTTGTATCATCAGTG 960  
961 |||||  
1018 GCAAGAAACGACACCGGATACATCACTGTTCTCTTCAAAGCATCCAGTCAATCAGCT 1077  
1019 |||||  
961 GCAAGAAACGACACCGGATACATCACTGTTCTCTTCAAAGCATCCAGTCAATCAGCT 1020  
1021 |||||  
1078 TTGGTTACCATCGTAGGAAAGGGAATTTAAATGTACCAATTCAGAGTGAAGATTATGAA 1137  
1079 |||||  
1021 TTGGTTACCATCGTAGGAAAGGGAATTTAAATGTACCAATTCAGAGTGAAGATTATGAA 1080  
1081 |||||  
1138 ATTGACCAATATGAAGAGTTTGTGTTTCTGTCAGGTTTAAAGCTTACCCCAATTCAGA 1197  
1139 |||||  
1081 ATTGACCAATATGAAGAGTTTGTGTTTCTGTCAGGTTTAAAGCTTACCCCAATTCAGA 1140  
1141 |||||  
1198 TGTACGTGGACCTTCTCTCGAAATCAATTTCTTGTGTGAGCAAAAGGGTCTTGATTAACGGA 1257

1141 TGTACCTGGACCTTCTCTCGAAATCAATTTCTTGTGAGCAAAAGGGTCTTGATTAACGGA 1200  
1142 |||||  
1258 TACAGCATATCCAAAGTTTGTGAATCAATGACCAAGCCAGGAGAAATATATTTCCATGCA 1317  
1259 |||||  
1201 TACAGCATATCCAAAGTTTGTGAATCAATGACCAAGCCAGGAGAAATATATTTCCATGCA 1260  
1202 |||||  
1318 GAAATGATGATGCCCAATTTTACCAAAATGTTTACCGTGAATATTAAGAAAGGAAACCTCAA 1377  
1319 |||||  
1261 GAAATGATGATGCCCAATTTTACCAAAATGTTTACCGTGAATATTAAGAAAGGAAACCTCAA 1320  
1262 |||||  
1378 GTGCTCGCAGAGCATCGGCAAGTCCAGCGTCTCTGTTTCTCGATGGATATCCCATTTACCA 1437  
1379 |||||  
1321 GTGCTCGCAGAGCATCGGCAAGTCCAGCGTCTCTGTTTCTCGATGGATATCCCATTTACCA 1380  
1322 |||||  
1438 TCTTGGACCTTGGAAATGAGTGTTCAGAAAGTCTCCCAA CTGCAAGAGAGATCACAGAA 1497  
1439 |||||  
1381 TCTTGGACCTTGGAAATGAGTGTTCAGAAAGTCTCCCAA CTGCAAGAGAGATCACAGAA 1440  
1441 |||||  
1498 GGAGTCTGGAAATGAAAGGCTTAAAGAAAGTGTTCGACAGTGGGTGTCAGAGTACT 1557  
1499 |||||  
1441 GGAGTCTGGAAATGAAAGGCTTAAAGAAAGTGTTCGACAGTGGGTGTCAGAGTACT 1500  
1501 |||||  
1558 CTAAACATGAGTGAAGCCATATAAAGGGTTCCTGTCAGAGTCTGTGATACAAATCCCTT 1617  
1559 |||||  
1501 CTAAACATGAGTGAAGCCATATAAAGGGTTCCTGTCAGAGTCTGTGATACAAATCCCTT 1560  
1502 |||||  
1618 GGACATCTTGTGAGAGTCTCTTTTAACTCTCCAGGCCCTTCCCTTTTCATCCAGAC 1677  
1619 |||||  
1561 GGACATCTTGTGAGAGTCTCTTTTAACTCTCCAGGCCCTTCCCTTTTCATCCAGAC 1620  
1621 |||||  
1678 AACATCTCTTATGCAACATTTGGTGTGTTGTCCTCTCTTCAATGTCGTTTAAACCTG 1737  
1679 |||||  
1621 AACATCTCTTATGCAACATTTGGTGTGTTGTCCTCTCTTCAATGTCGTTTAAACCTG 1680  
1622 |||||  
1738 CTAATTTGTCAACAAAGTCAAAAAGCAATTTAGGTATGAAAGCCAGCTTACAGATGATGAT 1797  
1739 |||||  
1681 CTAATTTGTCAACAAAGTCAAAAAGCAATTTAGGTATGAAAGCCAGCTTACAGATGATGAT 1740  
1741 |||||  
1798 GTGACCGGCTCTCAGATAATGATGATCTCTTAAGTGTGATTTTCAAGAAATATGATATGAT 1857  
1799 |||||  
1741 GTGACCGGCTCTCAGATAATGATGATCTCTTAAGTGTGATTTTCAAGAAATATGATATGAT 1800  
1801 |||||  
1858 CT-----CAATGGGAGTTTCCAAAGAGAAATTTA 1887  
1859 |||||  
1801 CTCAATGGGAGTTTCCAAAGAGAAATTTGCA CAAATGGGAGTTTCCAAAGAGAAATTTA 1860  
1802 |||||  
1888 GAGTTTGGGAGGTTACTAGGATCAGTGTCTTTTGGAAAGTGTGATGAAACGCAACAGCTTAT 1947  
1889 |||||  
1861 GAGTTTGGGAGGTTACTAGGATCAGTGTCTTTTGGAAAGTGTGATGAAACGCAACAGCTTAT 1920  
1862 |||||  
1948 GGAAATTAGCAAAACAGGAGTCTCAATCCAGGTTGCCGTCAAAATGCTGAAAGAAAGCA 2007  
1949 |||||  
1921 GGAAATTAGCAAAACAGGAGTCTCAATCCAGGTTGCCGTCAAAATGCTGAAAGAAAGCA 1980  
1922 |||||  
2008 GACAGCTCTGAAAGAGAGGCACTCATGTGAGAACTCAAGATGATGATGACCCAGCTGGGAAGC 2067  
2009 |||||  
1981 GACAGCTCTGAAAGAGAGGCACTCATGTGAGAACTCAAGATGATGATGACCCAGCTGGGAAGC 2040  
1982 |||||  
2068 CACGAGATTTGTGAA CCTGCTGGGGCGTGCACTGTGAGGACCAATTTACTTTGAT 2127  
2069 |||||  
2041 CACGAGATTTGTGAA CCTGCTGGGGCGTGCACTGTGAGGACCAATTTACTTTGAT 2100  
2101 |||||  
2128 TTTGAAATCTGTTGTAATGTTGATCTTCTCAACTATCTAAGAGTAAAGAGAAATTT 2187  
2129 |||||  
2101 TTTGAAATCTGTTGTAATGTTGATCTTCTCAACTATCTAAGAGTAAAGAGAAATTT 2160  
2102 |||||  
2188 CACAGGACTTGGACAGAGATTTTCAAGGAAACAAATTTTACGTTTACCCCACTTTCCAA 2247  
2189 |||||  
2161 CACAGGACTTGGACAGAGATTTTCAAGGAAACAAATTTTACGTTTACCCCACTTTCCAA 2220  
2162 |||||  
2248 TCACATCCAAATTCAGCATGCTGTTTCAAGAGTTCAGATACACCCGAGCTCGGAT 2307  
2249 |||||

Db	2221	TCACATCCAAATTCAGACATGCTGTTCAAGAGAAAGTTTCAGATACACCCGGACTCGGAT	2280
Qy	2308	CAAAATCTCAGGCTTCATGGGAATTCATTTCTCACTCTGAAGATGAATTCGAATATGAAAAC	2367
Db	2281	CAAAATCTCAGGCTTCATGGGAATTCATTTCTCACTCTGAAGATGAATTCGAATATGAAAAC	2340
Qy	2368	CAAAAAGGCTGGAAGAGAGAGAGGACTTTGAATGTGCTTTACATTTGAAGATCTTCTTTGC	2427
Db	2341	CAAAAAGGCTGGAAGAGAGAGAGGACTTTGAATGTGCTTTACATTTGAAGATCTTCTTTGC	2400
Qy	2428	TTTGCATATCAATGTCGCAAGGAATGGAATTTCTGGAATTTAAAGTCGTGTTCACAGA	2487
Db	2401	TTTGCATATCAATGTCGCAAGGAATGGAATTTCTGGAATTTAAAGTCGTGTTCACAGA	2460
Qy	2488	GACCTGGCCGCCAGGAACTGTCTTGTCAACCAGGAAAGTGTGAAGATATGTGACTTT	2547
Db	2461	GACCTGGCCGCCAGGAACTGTCTTGTCAACCAGGAAAGTGTGAAGATATGTGACTTT	2520
Qy	2548	GGATTGGCTCGAGATATCATAGATGATTCCAACTATGTTGTCAAGGGCAATGCCCGTCTG	2607
Db	2521	GGATTGGCTCGAGATATCATAGATGATTCCAACTATGTTGTCAAGGGCAATGCCCGTCTG	2580
Qy	2608	CCTGTAAATGATGGCCCCCGAAGCCCTGTTTGAAGGCATCTACACCATTAAGAGTGAT	2667
Db	2581	CCTGTAAATGATGGCCCCCGAAGCCCTGTTTGAAGGCATCTACACCATTAAGAGTGAT	2640
Qy	2668	GTCTGTCTATATGGAATATTTACTGTGGGAAATCTTCTCACTTGTGTGAATCCTTACCCT	2727
Db	2641	GTCTGTCTATATGGAATATTTACTGTGGGAAATCTTCTCACTTGTGTGAATCCTTACCCT	2700
Qy	2728	GGCATTCCGGTTGATGCTAACTTCTCAAACTGAATCAAAATGGATTTTAAATGGATCAG	2787
Db	2701	GGCATTCCGGTTGATGCTAACTTCTCAAACTGAATCAAAATGGATTTTAAATGGATCAG	2760
Qy	2788	CCATTTTATGCTACAGAGAAATATACATTAATGCAATCCTCTGGGCTTTTGACTCA	2847
Db	2761	CCATTTTATGCTACAGAGAAATATACATTAATGCAATCCTCTGGGCTTTTGACTCA	2820
Qy	2848	AGGAAACGGCCATCCTTCCCTAATTTGACTTCGTTTTTTAGGATGTCAAGCTGGCAGATGCA	2907
Db	2821	AGGAAACGGCCATCCTTCCCTAATTTGACTTCGTTTTTTAGGATGTCAAGCTGGCAGATGCA	2880
Qy	2908	GAAGAAAGCGATGTATCAGAAATGTGGATGGCCGTTGTTTCGGAATGTCTCTACACCTACCAA	2967
Db	2881	GAAGAAAGCGATGTATCAGAAATGTGGATGGCCGTTGTTTCGGAATGTCTCTACACCTACCAA	2940
Qy	2968	AACAGGCGA	2976
Db	2941	AACAGGCGA	2949

Search completed: August 27, 2003, 16:01:41  
Job time : 1052.15 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 28, 2003, 01:02:19 ; Search time 1021.5 Seconds  
(without alignments)  
7863.095 Million cell updates/sec

Title: US-09-919-408A-3

Perfect score: 3501

Sequence: 1 CGAGGGCGGATCCGAGGCT.....AAAAAAAAAAAAAAAAAAAA 3501

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1533700 seqs, 1147125425 residues

Total number of hits satisfying chosen parameters: 3067400

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3501	100.0	3501	9	US-09-919-408-3
2	3501	100.0	3501	10	US-09-872-136-3
3	3083.2	88.1	3120	10	US-09-982-610-22
4	2227.2	63.6	3453	9	US-09-919-408-1
5	2227.2	63.6	3453	10	US-09-872-136-1
6	633	18.1	645	10	US-09-796-692-7538
7	633	18.1	645	14	US-10-040-862-8085
8	608.4	17.4	610	14	US-09-796-692-8085
9	608.4	17.4	610	14	US-10-040-862-8085
10	606.4	17.3	608	10	US-09-796-692-8132
11	606.4	17.3	608	14	US-10-040-862-8132
12	605	17.3	605	10	US-09-796-692-8927
13	605	17.3	605	14	US-10-040-862-8927
14	598.4	17.1	601	10	US-09-796-692-8273
15	598.4	17.1	601	14	US-10-040-862-8273
16	597	17.1	597	10	US-09-796-692-7976

17	597	17.1	597	14	US-10-040-862-7976	Sequence 7976, Ap
18	595.4	17.0	597	10	US-09-796-692-9010	Sequence 9010, Ap
19	595.4	17.0	597	14	US-10-040-862-9010	Sequence 9010, Ap
20	593.2	16.9	598	10	US-09-796-692-8566	Sequence 8566, Ap
21	593.2	16.9	598	14	US-10-040-862-8566	Sequence 8566, Ap
22	590.8	16.9	594	10	US-09-796-692-8135	Sequence 8135, Ap
23	590.8	16.9	594	14	US-10-040-862-8135	Sequence 8135, Ap
24	590.4	16.9	605	10	US-09-796-692-8047	Sequence 8047, Ap
25	590.4	16.9	605	14	US-10-040-862-8047	Sequence 8047, Ap
26	583.8	16.7	589	10	US-09-796-692-8630	Sequence 8630, Ap
27	583.8	16.7	589	14	US-10-040-862-8630	Sequence 8630, Ap
28	562.4	16.1	564	10	US-09-796-692-9025	Sequence 9025, Ap
29	562.4	16.1	564	14	US-10-040-862-9025	Sequence 9025, Ap
30	554.4	15.8	556	10	US-09-796-692-7883	Sequence 7883, Ap
31	554.4	15.8	556	14	US-10-040-862-7883	Sequence 7883, Ap
32	541	15.5	550	10	US-09-796-692-8494	Sequence 8494, Ap
33	541	15.5	550	14	US-10-040-862-8494	Sequence 8494, Ap
34	539	15.4	539	10	US-09-796-692-8255	Sequence 8255, Ap
35	539	15.4	539	14	US-10-040-862-8255	Sequence 8255, Ap
36	510	14.6	510	10	US-09-796-692-8268	Sequence 8268, Ap
37	510	14.6	510	14	US-10-040-862-8268	Sequence 8268, Ap
38	502.8	14.4	508	10	US-09-796-692-8032	Sequence 8032, Ap
39	502.8	14.4	508	14	US-10-040-862-8032	Sequence 8032, Ap
40	491.4	14.0	497	10	US-09-796-692-8408	Sequence 8408, Ap
41	491.4	14.0	497	14	US-10-040-862-8408	Sequence 8408, Ap
42	478.4	13.7	606	10	US-09-796-692-8114	Sequence 8114, Ap
43	478.4	13.7	606	14	US-10-040-862-8114	Sequence 8114, Ap
44	446.4	12.8	581	10	US-09-796-692-9031	Sequence 9031, Ap
45	446.4	12.8	581	14	US-10-040-862-9031	Sequence 9031, Ap

#### ALIGNMENTS

#### RESULT 1

US-09-919-408-3  
; Sequence 3, Application US/09919408  
; Patent No. US20020072077A1  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Imclone Systems Incorporated  
; STREET: 180 Varick Street  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentcin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/919,408  
; FILING DATE: 31-Jul-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/977,451  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: US 07/906,397  
; FILING DATE: 26-JUN-1992  
; APPLICATION NUMBER: US PCT/US92/05401  
; FILING DATE: 26-JUN-1992  
; APPLICATION NUMBER: TW 81102961  
; FILING DATE: 15-APR-1992  
; APPLICATION NUMBER: US PCT/US92/02750  
; FILING DATE: 02-APR-1992  
; APPLICATION NUMBER: US 07/813,593  
; FILING DATE: 24-DEC-1991  
; APPLICATION NUMBER: US 07/793,065



FILING DATE: 15-NOV-1991  
APPLICATION NUMBER: US 07/728,913  
FILING DATE: 28-JUN-1991  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Feit, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3501 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
FEATURE:

NAME/KEY: CDS  
LOCATION: 58..3039

FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 139..3036

FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 58..138

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-919-408-3

Query Match 100.0%; Score 3501; DB 9; Length 3501;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGAGCGGCATCCGAGGGCTGGCGCGCCCTGGGGGACCCCGGGCTCCGAGGCCATG 60  
Db 1 CGAGCGGCATCCGAGGGCTGGCGCGCCCTGGGGGACCCCGGGCTCCGAGGCCATG 60

Qy 61 CGCGGCTGGCGCGGACCGGGGACCGTGGCGTCTGTTGTTTTCGCAATGATA 120  
Db 61 CGCGGCTGGCGCGGACCGGGGACCGTGGCGTCTGTTGTTTTCGCAATGATA 120

Qy 121 TTTGGGACTATTCAAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATCAAG 180  
Db 121 TTTGGGACTATTCAAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATCAAG 180

Qy 181 AACAAATGATTTCAGTGGGGAAGTCATCATATCCCATGTTATCAGAAATCCCGGAA 240  
Db 181 AACAAATGATTTCAGTGGGGAAGTCATCATATCCCATGTTATCAGAAATCCCGGAA 240

Qy 241 GACCTCGGGTGTGGTTGAGACCCAGAGCTCAGGGACAGTGTACGAGCTGCCGTGTG 300  
Db 241 GACCTCGGGTGTGGTTGAGACCCAGAGCTCAGGGACAGTGTACGAGCTGCCGTGTG 300

Qy 301 GAAGTGAATGTCTGCTTCCATCACTGCAAGTGTGTCGATGCCAGGGGAACATT 360  
Db 301 GAAGTGAATGTCTGCTTCCATCACTGCAAGTGTGTCGATGCCAGGGGAACATT 360

Qy 361 TCCTGTCTCGGGTCTTTAAGCACAGCTCCCTGAATTCAGCCACATTTTGAATTTACAA 420  
Db 361 TCCTGTCTCGGGTCTTTAAGCACAGCTCCCTGAATTCAGCCACATTTTGAATTTACAA 420

Qy 421 AACAGAGAGTGTTCATGGTCAATTTGAAATGACAGAAACCCAGCTGGAGATAC 480  
Db 421 AACAGAGAGTGTTCATGGTCAATTTGAAATGACAGAAACCCAGCTGGAGATAC 480

Qy 481 CTACTTTTATTACAGTGAAGCTACCAATTTACCAATTTTGTGTACAGTGAATGA 540  
Db 481 CTACTTTTATTACAGTGAAGCTACCAATTTACCAATTTTGTGTACAGTGAATGA 540

541 AATACCTCTGCTTTACACATTAAGAGACCTTACTTTAGAAAAATGAAAAACGAGCGCC 600  
541 AATACCTCTGCTTTACACATTAAGAGACCTTACTTTAGAAAAATGAAAAACGAGCGCC 600  
601 CTGCTCTGCATATCTGAGAGCGTTTCAGAGCCCATCGTGAATGGGTGCTTTGCGATTCA 660  
601 CTGCTCTGCATATCTGAGAGCGTTTCAGAGCCCATCGTGAATGGGTGCTTTGCGATTCA 660  
661 CAGGGGGAAGCTGTAAAGAGAAAGTCCAGCTGTTTAAAGAGGAGGAAAAAGTGCCTT 720  
661 CAGGGGGAAGCTGTAAAGAGAAAGTCCAGCTGTTTAAAGAGGAGGAAAAAGTGCCTT 720  
721 CATGAATTTATTTGGGACGGACATAAGGTGTCGTGCCAGAAATGAACCTGGGAGGAATGC 780  
721 CATGAATTTATTTGGGACGGACATAAGGTGTCGTGCCAGAAATGAACCTGGGAGGAATGC 780  
781 ACCAGGCTGTTTCAATATAGATCTAAATCAAACTCCTCAGACCACATTCCTCAATTTATTT 840  
781 ACCAGGCTGTTTCAATATAGATCTAAATCAAACTCCTCAGACCACATTCCTCAATTTATTT 840  
841 CTTAAAGTAGGGGAACCCCTTTATGGATAAGGTGCAAAAGCTGTTTCAATGTGAACCATGATTC 900  
841 CTTAAAGTAGGGGAACCCCTTTATGGATAAGGTGCAAAAGCTGTTTCAATGTGAACCATGATTC 900  
901 GGGCTCACCTGGGAAATTAGAAAAAAGGCACTCGAGGAGGGCAACTACTTTTGAGATGAGT 960  
901 GGGCTCACCTGGGAAATTAGAAAAAAGGCACTCGAGGAGGGCAACTACTTTTGAGATGAGT 960  
961 ACCTATTTCAACAAACAGAACTATGATACGGATTCGTTTCTTTTGTGTATCATCAGTGGCA 1020  
961 ACCTATTTCAACAAACAGAACTATGATACGGATTCGTTTCTTTTGTGTATCATCAGTGGCA 1020  
1021 AGAAACGACACCGGATACTACACTTCTTCAAGAGTCCAGTCAATCAGCTTTG 1080  
1021 AGAAACGACACCGGATACTACACTTCTTCAAGAGTCCAGTCAATCAGCTTTG 1080  
1081 GTTACCATCTAGAGGAGGATTAATGCTACCAATTTCAAGTGAAGATTAATGAAATTT 1140  
1081 GTTACCATCTAGAGGAGGATTAATGCTACCAATTTCAAGTGAAGATTAATGAAATTT 1140  
1141 GACCAATATCAAGAGTCTTCTGTCAGGTTTAAAGCCTACCCACAAATCAGATGT 1200  
1141 GACCAATATCAAGAGTCTTCTGTCAGGTTTAAAGCCTACCCACAAATCAGATGT 1200  
1201 ACCTGGACCTTCTCTCGAAATCATTTCTTGTGAGCAAAAGGCTCTTTGATAACGGATAC 1260  
1201 ACCTGGACCTTCTCTCGAAATCATTTCTTGTGAGCAAAAGGCTCTTTGATAACGGATAC 1260  
1261 AGCATATCCAAGTTTTCGAAATCAAGCAACAGCCAGGAGAAATATATATTCATGCGAGAA 1320  
1261 AGCATATCCAAGTTTTCGAAATCAAGCAACAGCCAGGAGAAATATATATTCATGCGAGAA 1320  
1321 AATGATGATCCCAATTTTCCAAATGTTCCGCTGAATTAAGAGGAAACCTCAAGTG 1380  
1321 AATGATGATCCCAATTTTCCAAATGTTCCGCTGAATTAAGAGGAAACCTCAAGTG 1380  
1381 CTGCGAAGCATCGGCAAGTCTGAGGCTCTGTTTCTCGATGGATACCCATTTACCATCT 1440  
1381 CTGCGAAGCATCGGCAAGTCTGAGGCTCTGTTTCTCGATGGATACCCATTTACCATCT 1440  
1441 TGGACCTCGGAGAGTGTTCAGCAAGTCTCCAACTGCGACAGAGAGATCACAAGAGGA 1500  
1441 TGGACCTCGGAGAGTGTTCAGCAAGTCTCCAACTGCGACAGAGAGATCACAAGAGGA 1500  
1501 GTCTGGAATAGAAAGGCTTAAAGAGTGTTCGACAGTGGGTGTCGAGCAGTACTCTA 1560  
1501 GTCTGGAATAGAAAGGCTTAAAGAGTGTTCGACAGTGGGTGTCGAGCAGTACTCTA 1560  
1561 AACATGATGAGGCAATAAAGGGTTCCTGCTCAAGTGTGTGATACAAATTCCTCTGGC 1620  
1561 AACATGATGAGGCAATAAAGGGTTCCTGCTCAAGTGTGTGATACAAATTCCTCTGGC 1620

1621 ACATCTTGTCAGAGATCCCTTTTAAACTCTCCAGGCCCTTCCCTTTCATCCAGACAAC 1680  
1621 ACATCTTGTCAGAGATCCCTTTTAAACTCTCCAGGCCCTTCCCTTTCATCCAGACAAC 1680  
1681 ATCTCATCTATGCAACAAATGGTGTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740  
1681 ATCTCATCTATGCAACAAATGGTGTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740  
1741 ATTTGTCAACAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
1741 ATTTGTCAACAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
1801 ACCGGCTCTCCAGATATGATCTCTACGTGTCATTTTCAAGATTTTCAAGATTTTCAAGATTT 1860  
1801 ACCGGCTCTCCAGATATGATCTCTACGTGTCATTTTCAAGATTTTCAAGATTTTCAAGATTT 1860  
1861 AAATGGAGTTTCCAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920  
1861 AAATGGAGTTTCCAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920  
1921 GGAAGTGTATGAAACGCAACAGCTTATGGAATTTAGCAAAACAGAGGTCTCAATCCAGGTT 1980  
1921 GGAAGTGTATGAAACGCAACAGCTTATGGAATTTAGCAAAACAGAGGTCTCAATCCAGGTT 1980  
1981 GCCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATCTCAGAA 2040  
1981 GCCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATCTCAGAA 2040  
2041 CTCAAGATGATGACCCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGGCGTGC 2100  
2041 CTCAAGATGATGACCCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGGCGTGC 2100  
2101 ACATGTCAGGACCAATTTACTTGTATTTTGAATCTGTGTCATGTCATCTCTCAAC 2160  
2101 ACATGTCAGGACCAATTTACTTGTATTTTGAATCTGTGTCATGTCATCTCTCAAC 2160  
2161 TATCTAAGATGAAAGAGAAAATTTTCAAGGATTTTGAAGATTTTCAAGGAAACAC 2220  
2161 TATCTAAGATGAAAGAGAAAATTTTCAAGGATTTTGAAGATTTTCAAGGAAACAC 2220  
2221 AATTTCAAGTTTACCCCACTTTTCCAAATCATCTCCAAATTTCCAGATGCTGTGTTCAAGA 2280  
2221 AATTTCAAGTTTACCCCACTTTTCCAAATCATCTCCAAATTTCCAGATGCTGTGTTCAAGA 2280  
2281 GAAATTCAGATACACCCGAGCTCGGATCAAAATCTCAGGGCTTCATGGAATTTCAATTTCAAC 2340  
2281 GAAATTCAGATACACCCGAGCTCGGATCAAAATCTCAGGGCTTCATGGAATTTCAATTTCAAC 2340  
2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400  
2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400  
2401 GTCTTACATTTGAGATCTTTTGTGTCATATCAAGTTCGCAAGGATTTGGAATTT 2460  
2401 GTCTTACATTTGAGATCTTTTGTGTCATATCAAGTTCGCAAGGATTTGGAATTT 2460  
2461 CTGGAATTTAAGTCTGTGTTTCAAGAGACTGCGCCAGGAACTGCTGTGTCACCCAC 2520  
2461 CTGGAATTTAAGTCTGTGTTTCAAGAGACTGCGCCAGGAACTGCTGTGTCACCCAC 2520  
2521 GGAAGTGTGGAAGATATGATGATTTGGAATTTGGAATTTGGAATTTGGAATTTGGAATTT 2580  
2521 GGAAGTGTGGAAGATATGATGATTTGGAATTTGGAATTTGGAATTTGGAATTTGGAATTT 2580  
2581 TATGTTGTCAGGGGCAATCCCGTCTGCTGTAATGGAATGGAATGGAATGGAATGGAATGGAAT 2640  
2581 TATGTTGTCAGGGGCAATCCCGTCTGCTGTAATGGAATGGAATGGAATGGAATGGAATGGAAT 2640  
2641 GAAGGATCTACACCAATTAAGAGTATGTCGTCATATGGAATTTTACTGTGGGAATC 2700  
2641 GAAGGATCTACACCAATTAAGAGTATGTCGTCATATGGAATTTTACTGTGGGAATC 2700  
2701 TTCTCACTGGTGTGAATCCTTACCTGGCAATTCGGGTGTGATGCTTAATCTTCTACAAACTG 2760

2701 TTCTCACTGGTGTGAATCCTTACCTGGCAATTCGGGTGTGATGCTTAATCTTCTACAACTG 2760  
2761 ATTCAAAATGGATTTAAATGGATCAGCCATTTTATGCTACAGAAAGAAATATACATATATA 2820  
2761 ATTCAAAATGGATTTAAATGGATCAGCCATTTTATGCTACAGAAAGAAATATACATATATA 2820  
2821 ATGCAATCCCTGCTGGGCTTTTGAATCAAGGAAACGGCCATCTTCCCTTAATTTGACTTCG 2880  
2821 ATGCAATCCCTGCTGGGCTTTTGAATCAAGGAAACGGCCATCTTCCCTTAATTTGACTTCG 2880  
2881 TTTTGTAGATGTCAGCTGGCAGATGCAAGAAAGCCAGATGATCAGAAATGTCGATGGCCGT 2940  
2881 TTTTGTAGATGTCAGCTGGCAGATGCAAGAAAGCCAGATGATCAGAAATGTCGATGGCCGT 2940  
2941 GTTTCGGAATGTCCTCACTACCTACCAAAACAGGCGACTTTTACGACAGAGATGGATTTG 3000  
2941 GTTTCGGAATGTCCTCACTACCTACCAAAACAGGCGACTTTTACGACAGAGATGGATTTG 3000  
3001 GGGTACTCTCTCCGAGGCTCAGGTCGAAATTCGTAGAGGAACTTTTGTAGTTTAAAG 3060  
3001 GGGTACTCTCTCCGAGGCTCAGGTCGAAATTCGTAGAGGAACTTTTGTAGTTTAAAG 3060  
3061 ACTTTCATCCCTCCACCTATCCCTAAACAGGCTGTAGATTTACCAAAACAGATTTAATTTTCA 3120  
3061 ACTTTCATCCCTCCACCTATCCCTAAACAGGCTGTAGATTTACCAAAACAGATTTAATTTTCA 3120  
3121 CACTTAAAGAAAATCTATTTATCAACTGCTGCTTCAACAGACTTTTCTCTAGAAGCCGTCT 3180  
3121 CACTTAAAGAAAATCTATTTATCAACTGCTGCTTCAACAGACTTTTCTCTAGAAGCCGTCT 3180  
3181 GCGTTTACTCTGTTTTTCAAAAGGCACTTTTGTAAATCAAAATCATCTCTGTCACAAAGCAG 3240  
3181 GCGTTTACTCTGTTTTTCAAAAGGCACTTTTGTAAATCAAAATCATCTCTGTCACAAAGCAG 3240  
3241 GAGGCTGATATGATGATTTTATTTGGAGCATGATCTGCAATCCAGGCTTCTCAGGCGG 3300  
3241 GAGGCTGATATGATGATTTTATTTGGAGCATGATCTGCAATCCAGGCTTCTCAGGCGG 3300  
3301 GCTTGTAGTGAATTTGTGTACCTGAAAGTACAGTATATTTCTGTAAATACATAAAACAAAGC 3360  
3301 GCTTGTAGTGAATTTGTGTACCTGAAAGTACAGTATATTTCTGTAAATACATAAAACAAAGC 3360  
3361 ATTTTGTAGAGGAAAGCTAAATATGATTTTAAAGTCTATGTTTTTAAATTAATATGTAAA 3420  
3361 ATTTTGTAGAGGAAAGCTAAATATGATTTTAAAGTCTATGTTTTTAAATTAATATGTAAA 3420  
3421 TTTTTCAGCTATTTAGTATATATTTTATGGTGGGAATAAAATTTCTCTACAGAAAAA 3480  
3421 TTTTTCAGCTATTTAGTATATATTTTATGGTGGGAATAAAATTTCTCTACAGAAAAA 3480  
3481 AAAAAAAAAAAAAAAAAAAAAA 3501  
3481 AAAAAAAAAAAAAAAAAAAAAA 3501

## RESULT 2

US-09-872-136-3

; Sequence 3, Application US/09872136

; Patent No. US20020119545A1

; GENERAL INFORMATION:

; APPLICANT: Lemischka, Ihor R.

; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL

; RECEPTORS AND THEIR LIGANDS

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESS: ImClone Systems Incorporated

; STREET: 180 Varick Street

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10014

; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/872,136  
FILING DATE: 01-Jun-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/208,786  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US/09/021,324  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US/07/977,451  
FILING DATE: 1992-11-19

APPLICATION NUMBER: US 07/906,397  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: US PCT/US92/05401  
FILING DATE: 26-JUN-1992

APPLICATION NUMBER: TW 81102961  
FILING DATE: 15-APR-1992  
APPLICATION NUMBER: US PCT/US92/02750  
FILING DATE: 02-APR-1992

APPLICATION NUMBER: US 07/813,593  
FILING DATE: 24-DEC-1991  
APPLICATION NUMBER: US 07/793,065  
FILING DATE: 15-NOV-1991

APPLICATION NUMBER: US 07/728,913  
FILING DATE: 28-JUN-1991  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991

ATTORNEY/AGENT INFORMATION:  
NAME: Feit, Irving N.  
REGISTRATION NUMBER: 28,601

REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054

INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3501 base pairs  
TYPE: nucleic acid

STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO

ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
FEATURE:

NAME/KEY: CDS  
LOCATION: 58..3039  
FEATURE:

NAME/KEY: mat.peptide  
LOCATION: 139..3036  
FEATURE:

NAME/KEY: sig.peptide  
LOCATION: 58..138  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Query Match 100.0%; Score 3501; DB 10; Length 3501;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CGAGCGGCATCCGAGGCTGGCGCGCCCTGGGGACCCCGGGCTCCGGAGCCATG	60
Db	1	CGAGCGGCATCCGAGGCTGGCGCGCCCTGGGGACCCCGGGCTCCGGAGCCATG	60
Qy	61	CCGGCGTTGGCGCGGACCGGGACCGTCCGCTGCTGTTTTTTCTGCAATGATA	120
Db	61	CCGGCGTTGGCGCGGACCGGGACCGTCCGCTGCTGTTTTTTCTGCAATGATA	120
Qy	121	TTTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTTTAATCAATCAATAG	180

Db 1201 ACGTGACCTTCTCGAAAAATCAATTTCTGTGTGAGCAAAAGGGCTCTTGATAACGGATAC 1260  
Qy 1261 AGCATATCCAAAGTTTGTCAATCATATAGCAACAGCCAGGAGAAATATATATTTCCATGCGAA 1320  
Db 1261 AGCATATCCAAAGTTTGTCAATCATATAGCAACAGCCAGGAGAAATATATATTTCCATGCGAA 1320  
Qy 1321 AATGATGATGCCCAATTTACCAAAAATGTTCCAGCTGAATATAAGAGGAAACCTCAAGTG 1380  
Db 1321 AATGATGATGCCCAATTTACCAAAAATGTTCCAGCTGAATATAAGAGGAAACCTCAAGTG 1380  
Qy 1381 CTGCGAAGCATCGGCAAGTCAAGGCTCTGTTTCTCGGATGGATACCAATTAACCATCT 1440  
Db 1381 CTGCGAAGCATCGGCAAGTCAAGGCTCTGTTTCTCGGATGGATACCAATTAACCATCT 1440  
Qy 1441 TGGACCTGGAAGAGTGTTCAGACAAGTCTCCMACTGCAACAGAGATCACAGAAGGA 1500  
Db 1441 TGGACCTGGAAGAGTGTTCAGACAAGTCTCCMACTGCAACAGAGATCACAGAAGGA 1500  
Qy 1501 GTCGTGGAATGAAGGCTAACAGAAAAGTGTGGAAGTGGGTGTCGACGAGTACTCTA 1560  
Db 1501 GTCGTGGAATGAAGGCTAACAGAAAAGTGTGGAAGTGGGTGTCGACGAGTACTCTA 1560  
Qy 1561 AACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTCTGATCAATTCCTTTGGC 1620  
Db 1561 AACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTCTGATCAATTCCTTTGGC 1620  
Qy 1621 ACATCTGTGAGAGTCCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAAGAAC 1680  
Db 1621 ACATCTGTGAGAGTCCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAAGAAC 1680  
Qy 1681 ATCTCAATCTATGCAACAATGGTGTGCTCTCTCTTCAATGTGCTGTTTAAACCTGCTA 1740  
Db 1681 ATCTCAATCTATGCAACAATGGTGTGCTCTCTCTTCAATGTGCTGTTTAAACCTGCTA 1740  
Qy 1741 ATTTGTCAAGTACAAAAGCAATTTAGTGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
Db 1741 ATTTGTCAAGTACAAAAGCAATTTAGTGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
Qy 1801 ACCGGTCTCCAGATAATAGTACTTCTACGTTGATTTACAGAAATGAATATGATCTC 1860  
Db 1801 ACCGGTCTCCAGATAATAGTACTTCTACGTTGATTTACAGAAATGAATATGATCTC 1860  
Qy 1861 AATGGGAGTTTCAAGAGAAAATTTAGAGTTTGGGAAGGTACTAGATCAGGTGCTTTT 1920  
Db 1861 AATGGGAGTTTCAAGAGAAAATTTAGAGTTTGGGAAGGTACTAGATCAGGTGCTTTT 1920  
Qy 1921 GGAAGAGTATGAACGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
Db 1921 GGAAGAGTATGAACGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
Qy 1981 GCCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCAAGAA 2040  
Db 1981 GCCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCAAGAA 2040  
Qy 2041 CTCAGATGATGACCCAGCTGGGAAGCCACAGAAATTTGTGAACCTGTGGGGGCGTGC 2100  
Db 2041 CTCAGATGATGACCCAGCTGGGAAGCCACAGAAATTTGTGAACCTGTGGGGGCGTGC 2100  
Qy 2101 ACATGTGAGGACCAATTTACTGATTTTGAATCTGTTGCTATGTTGATCTCTCAAC 2160  
Db 2101 ACATGTGAGGACCAATTTACTGATTTTGAATCTGTTGCTATGTTGATCTCTCAAC 2160  
Qy 2161 TATCTAAGAGTAAAGAGAAAATTTACAGGACTTTGGACAGAGATTTTCAAGGAAACAC 2220  
Db 2161 TATCTAAGAGTAAAGAGAAAATTTACAGGACTTTGGACAGAGATTTTCAAGGAAACAC 2220  
Qy 2221 AATTTCAAGTTTATACCCCACTTTCCAAATCAATCCAAATTCACAGATGCTGTTCAAGA 2280  
Db 2221 AATTTCAAGTTTATACCCCACTTTCCAAATCAATCCAAATTCACAGATGCTGTTCAAGA 2280  
Qy 2281 GAAGTTCAGATACACCCGAGCTCGGATCAATCTCAGGCTTCATGGGAATTCATTTTCA 2340  
Db 2281 GAAGTTTCAAGATACACCCGAGCTCGGATCAATCTCAGGCTTCATGGGAATTCATTTTCA 2340

Qy 2341 TCTGAAGATGAATTTCAATATATGAAACCAAAAAAGGCTGGAAGAGGAGGACTTGAAT 2400  
Db 2341 TCTGAAGATGAATTTCAATATATGAAACCAAAAAAGGCTGGAAGAGGAGGACTTGAAT 2400  
Qy 2401 GTCTTTACATTTCAAGATCTTTCTTTGCTTTGCTATATCAAGTTGCTCAAAAGGAATGGAATTT 2460  
Db 2401 GTCTTTACATTTCAAGATCTTTCTTTGCTTTGCTATATCAAGTTGCTCAAAAGGAATGGAATTT 2460  
Qy 2461 CTGGAATTTAAGTCTGTGTGTTTCAAGAGCTGGCCGCGAGGACGCTGCTGTCAACCCAC 2520  
Db 2461 CTGGAATTTAAGTCTGTGTGTTTCAAGAGCTGGCCGCGAGGACGCTGCTGTCAACCCAC 2520  
Qy 2521 GCGAAGTGTGGAAGATATGTAAGTCTTTGGATTTGGCTCGAGATATCATGAGTATTCGAAC 2580  
Db 2521 GCGAAGTGTGGAAGATATGTAAGTCTTTGGATTTGGCTCGAGATATCATGAGTATTCGAAC 2580  
Qy 2581 TATGTTGTTCAGGGGCAATGCCCGTCTGCTGTAAAAATGGATGGCCCCGGAAGGCTGTTT 2640  
Db 2581 TATGTTGTTCAGGGGCAATGCCCGTCTGCTGTAAAAATGGATGGCCCCGGAAGGCTGTTT 2640  
Qy 2641 GAAGGCATCTACACCAATTAAGAGTGTCTGCTCATATGGAATTTACTGTGGGAATC 2700  
Db 2641 GAAGGCATCTACACCAATTAAGAGTGTCTGCTCATATGGAATTTACTGTGGGAATC 2700  
Qy 2701 TTCTCACTGGTGTGAATCCTTACCTGGCATTCGGTTGATGCTAACTTCTCAAACTG 2760  
Db 2701 TTCTCACTGGTGTGAATCCTTACCTGGCATTCGGTTGATGCTAACTTCTCAAACTG 2760  
Qy 2761 ATTCAAAAATGGAATTTAAAAATGGATCAGCCATTTTATGCTACAGAAAGAAATATACATATA 2820  
Db 2761 ATTCAAAAATGGAATTTAAAAATGGATCAGCCATTTTATGCTACAGAAAGAAATATACATATA 2820  
Qy 2821 ATGCAATCTGCTGGGCTTTTGAATCAAGGAAACGGCCATCTCTCCCTAAATTTGACTTCG 2880  
Db 2821 ATGCAATCTGCTGGGCTTTTGAATCAAGGAAACGGCCATCTCTCCCTAAATTTGACTTCG 2880  
Qy 2881 TTTTGTAGGATGTCACTGGCAGATGCAAGAGAGCGATGATCAAGATGTTGGATGGCCGT 2940  
Db 2881 TTTTGTAGGATGTCACTGGCAGATGCAAGAGAGCGATGATCAAGATGTTGGATGGCCGT 2940  
Qy 2941 GTTTCGGAATGTCTCACACTTACAAACAGGCGACTTTTCAAGAGAGATGGAATTTG 3000  
Db 2941 GTTTCGGAATGTCTCACACTTACAAACAGGCGACTTTTCAAGAGAGATGGAATTTG 3000  
Qy 3001 GGGCTACTCTCTCCGAGGCTCAGGTGCAAGATTTCTAGAGGAAACAATTTAGTTTAAAG 3060  
Db 3001 GGGCTACTCTCTCCGAGGCTCAGGTGCAAGATTTCTAGAGGAAACAATTTAGTTTAAAG 3060  
Qy 3061 ACTTCATCTCCCTCACTATCCCTAAAGGCTGTAGATTAACAAACAGATTAATTTTCAAT 3120  
Db 3061 ACTTCATCTCCCTCACTATCCCTAAAGGCTGTAGATTAACAAACAGATTAATTTTCAAT 3120  
Qy 3121 CACTTAAAGAAATCTATTAATCAACTGCTCTCACCAGACTTTTCTCTAGAGCCGCT 3180  
Db 3121 CACTTAAAGAAATCTATTAATCAACTGCTCTCACCAGACTTTTCTCTAGAGCCGCT 3180  
Qy 3181 GCGTTTACTCTGTTTCAAGGAGCTTTTGTAAAAATCAAAATCAATCTCTGTCACAGGCGAG 3240  
Db 3181 GCGTTTACTCTGTTTCAAGGAGCTTTTGTAAAAATCAAAATCAATCTCTGTCACAGGCGAG 3240  
Qy 3241 GAGGAGCTGATAATGAACCTTTTATTTGGAGCATTTGATCTGCAATCCAGGCCCTTCTCAGGCCG 3300  
Db 3241 GAGGAGCTGATAATGAACCTTTTATTTGGAGCATTTGATCTGCAATCCAGGCCCTTCTCAGGCCG 3300  
Qy 3301 GCTTGTAGTGAATTTGTACTCTGAGTACAGTATATTTCTGTAAATATACATAAAACAAAGC 3360  
Db 3301 GCTTGTAGTGAATTTGTACTCTGAGTACAGTATATTTCTGTAAATATACATAAAACAAAGC 3360  
Qy 3361 ATTTTGTAGGAGAGCAATATGATTTTTTAAAGTCTATGTTTAAATATATATGTTAAA 3420  
Db 3361 ATTTTGTAGGAGAGCAATATGATTTTTTAAAGTCTATGTTTAAATATATATGTTAAA 3420

Qy 3421 TTTTTCAGCTATTAGTGATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAA 3480  
Db 3421 TTTTTCAGCTATTAGTGATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAA 3480  
Qy 3481 AAAAAAAAAAAAAAAAAA 3501  
Db 3481 AAAAAAAAAAAAAAAAAA 3501

## RESULT 3

US-09-982-610-22  
; Sequence 22, Application US/09982610  
; Patent No. US2002014620A1  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; Bennett, Brian D.  
; Goeddel, David  
; Lee, James M.  
; Matthews, William  
; Tsai, Siao Ping  
; Wood, William I.  
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080

## COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/982,610  
FILING DATE: 17-Oct-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/446,648  
FILING DATE: 1996-MAY-23  
APPLICATION NUMBER: 08/222616  
FILING DATE: 04-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 40,378  
REFERENCE/DOCKET NUMBER: P0821P3PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168

## INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:  
LENGTH: 3120 base pairs  
TYPE: Nucleic Acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear

SEQUENCE DESCRIPTION: SEQ ID NO: 22:

US-09-982-610-22

Query Match 88.1%; Score 3083.2; DB 10; Length 3120;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 3097; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 58 ATGCGCGGTGGCGCGGACCGGCGCTGCTGTTTCTTCTGCAATG 117  
Db 1 ATGAGAGCGTGGCGCGGACCGGCGCTGCTGTTTCTTCTGCAATG 60  
Qy 118 ATATTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATCAT 177  
Db 61 ATATTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATCAT 120

Qy 178 AAGAACAAATGATTTCATCAGTGGGGAAGTCATCATCATATCCATGGTATCAGAAATCCCCG 237  
Db 121 AAGAACAAATGATTTCATCAGTGGGGAAGTCATCATCATATCCATGGTATCAGAAATCCCCG 180  
Qy 238 GAAGACCTCGGGTGTGGTTGAGACCCAGAGCTCAGGGAAGTGTAGAGTGTGCGCT 297  
Db 181 GAAGACCTCGGGTGTGGTTGAGACCCAGAGCTCAGGGAAGTGTAGAGTGTGCGCT 240  
Qy 298 GTGGAAGTGTATGATCTGCTTCCATCACTGCAAGTGTGGTGTGATGCCCCAGGGAAC 357  
Db 241 GTGGAAGTGTATGATCTGCTTCCATCACTGCAAGTGTGGTGTGATGCCCCAGGGAAC 300  
Qy 358 ATTTCCTGCTCTGGGTCTTTAAGCAGCTCCCTGTAATGCCAGCCACATTTGATTTA 417  
Db 301 ATTTCCTGCTCTGGGTCTTTAAGCAGCTCCCTGTAATGCCAGCCACATTTGATTTA 360  
Qy 418 CAAACAGAGAGGTTGTTCCATGGTCAATTTGAAATGACAGAAACCCAGCTGGAGAA 477  
Db 361 CAAACAGAGAGGTTGTTCCATGGTCAATTTGAAATGACAGAAACCCAGCTGGAGAA 420  
Qy 478 TACCTACTTTTATTCAGAGTGAAGCTACCAATTAACAATATTTGTTTACAGTGAATATA 537  
Db 421 TACCTACTTTTATTCAGAGTGAAGCTACCAATTAACAATATTTGTTTACAGTGAATATA 480  
Qy 538 AGAAATACCTGCTTTTACACATTAAGAAGACCTTACTTTAGAAAAATGGAACACAGGAC 597  
Db 481 AGAAATACCTGCTTTTACACATTAAGAAGACCTTACTTTAGAAAAATGGAACACAGGAC 540  
Qy 598 GCCCTGCTGTCATATCTGAGAGCGTTCCAGAGCGGATCTGGGAATGGTGTCTTTGCCAT 657  
Db 541 GCCCTGCTGTCATATCTGAGAGCGTTCCAGAGCGGATCTGGGAATGGTGTCTTTGCCAT 600  
Qy 658 TCACAGGGGAAAGCTGTAAAGAGAAAGTCCAGCTGTGTGTTTAAAAAGAGGAAAAAGTG 717  
Db 601 TCACAGGGGAAAGCTGTAAAGAGAAAGTCCAGCTGTGTGTTTAAAAAGAGGAAAAAGTG 660  
Qy 718 CTTTCATGAATTTTGGGACGGAATAAAGTGTGTCGAGCAATGAATGAGGAGGGA 777  
Db 661 CTTTCATGAATTTTGGGACGGAATAAAGTGTGTCGAGCAATGAATGAGGAGGGA 720  
Qy 778 TGCAACAGGCTGTTTCAACAATAGATCTAAATCAAACTCTCAGACCAATGCCCAAAATTA 837  
Db 721 TGCAACAGGCTGTTTCAACAATAGATCTAAATCAAACTCTCAGACCAATGCCCAAAATTA 780  
Qy 838 TTTCTTAAAGTAGGGGAACCTTATGGATAAGGTGCAAAAGCTGTTTCATGTGAACCATGGA 897  
Db 781 TTTCTTAAAGTAGGGGAACCTTATGGATAAGGTGCAAAAGCTGTTTCATGTGAACCATGGA 840  
Qy 898 TTGCGGCTCACCTGGGAATTAAGAAAACAAGCACTCGAGGAGGGCAACTCTTTGAGATG 957  
Db 841 TTGCGGCTCACCTGGGAATTAAGAAAACAAGCACTCGAGGAGGGCAACTCTTTGAGATG 900  
Qy 958 AGTACCTATTCAACAAACAGAACTATGATCAGGATCTGTTGTTTGTGTTTGTATCATCAGTG 1017  
Db 901 AGTACCTATTCAACAAACAGAACTATGATCAGGATCTGTTGTTTGTGTTTGTATCATCAGTG 960  
Qy 1018 GCAAGAAAACGACACCGGATACATCACTGTTTCTCTTCAAAGCATCCAGTCAATCAGCT 1077  
Db 961 GCAAGAAAACGACACCGGATACATCACTGTTTCTCTTCAAAGCATCCAGTCAATCAGCT 1020  
Qy 1078 TTGTTTACCATCGTAGGAAAGGGATTTTATAAATGCTACCAATTCAGTGAAGATTAATGAA 1137  
Db 1021 TTGTTTACCATCGTAGGAAAGGGATTTTATAAATGCTACCAATTCAGTGAAGATTAATGAA 1080  
Qy 1138 ATTGACCAATATGAAGAGTTTGTGTTTCTGTCAGGTTTAAAGCTACCCCAAAATCAGA 1197  
Db 1081 ATTGACCAATATGAAGAGTTTGTGTTTCTGTCAGGTTTAAAGCTACCCCAAAATCAGA 1140  
Qy 1198 TGTACGTGGACCTTCTCTCGAAATCATTTCTGTCAGCAAAAGGGTCTTTGTAACCGGA 1257  
Db 1141 TGTACGTGGACCTTCTCTCGAAATCATTTCTGTCAGCAAAAGGGTCTTTGTAACCGGA 1200  
Qy 1258 TACAGCATATCCAAGTTTTTGCATATATAGCACCGCAGGAGAAATATATATTCATGCA 1317

1201 TACAGCATATCCAGTTTGGCAATCATAGACACAGCCGAGGAAATATATATTCATGCA 1260  
1318 GAAATGATGATGCCCAATTTACAAATGTTTCAGCTGAAATATAAGAGGAAACCTCAA 1377  
1261 GAAATGATGATGCCCAATTTACCAAAATGTTTCAGCTGTAATATAAGAGGAAACCTCAA 1320  
1378 GTCTCGCAGAGCATCGGCAAGTCAGGCGTCTGTTTCTCGGATGGATCCCATTAACA 1437  
1321 GTCTCGCAGAGCATCGGCAAGTCAGGCGTCTGTTTCTCGGATGGATCCCATTAACA 1380  
1438 TCTTGGACCTCGGAAGAGTGTTCAGACAAAGTCTCCCAACTGCACAGAAGAGATCACAGAA 1497  
1381 TCTTGGACCTCGGAAGAGTGTTCAGACAAAGTCTCCCAACTGCACAGAAGAGATCACAGAA 1440  
1498 GGAGTCTGGAAATAGAAAGGCTAAACAGAAAGTGTTCAGACAGTGGGTGTCAGAGTACT 1557  
1441 GGAGTCTGGAAATAGAAAGGCTAAACAGAAAGTGTTCAGACAGTGGGTGTCAGAGTACT 1500  
1558 CTAACATGATGAGTGAAGCATAAAGAGTTCCTGCTCAAGTCTGTGCATACAAATTCCTTT 1617  
1501 CTAACATGATGAGTGAAGCATAAAGAGTTCCTGCTCAAGTCTGTGCATACAAATTCCTTT 1560  
1618 GGACATCTTGTGAGACGATCCTTTAAACTCTCCAGGCCCTTCCCTTTTCATCAAGAC 1677  
1561 GGACATCTTGTGAGACGATCCTTTAAACTCTCCAGGCCCTTCCCTTTTCATCAAGAC 1620  
1678 AACATCTCAATCTATGCAACAAATTTGGTGTTCCTCTCTTCAATGTGCTTTTAAACCTG 1737  
1621 AACATCTCAATCTATGCAACAAATTTGGTGTTCCTCTCTTCAATGTGCTTTTAAACCTG 1680  
1738 CTAATTTGTCAACAGTCAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAG 1797  
1681 CTAATTTGTCAACAGTCAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAG 1740  
1798 GTGACCGGCTCCTCAGATATAGTACTCTTCACTGTTGATTTTTCAGAGATATGAATATGAT 1857  
1741 GTGACCGGATCCTCAGATATAGTACTCTTCACTGTTGATTTTTCAGAGATATGAATATGAT 1800  
1858 CTCAATTTGGAGTTTCCAAAGAGAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCT 1917  
1801 GTCAATTTGGAGTTTCCAAAGAGAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCT 1860  
1918 TTTGGAAAGTGTAAACGCAACAGTGTATGGAATTTAGCAAAAACAGGAGTCTCAATCCAG 1977  
1861 TTTGGAAAGTGTAAACGCAACAGTGTATGGAATTTAGCAAAAACAGGAGTCTCAATCCAG 1920  
1978 GTTGGCTCAAAATGCTGAAAGAAAGAGACAGACAGCTCTGAAAGAGGCACTCATGTCA 2037  
1921 GTTACCGTCAAAATGCTGAAAGAAAGAGACAGACAGCTCTGAAAGAGGCACTCATGTCA 1980  
2038 GAACTCAAGATGATGACCCAGCTGGGAAGCCACAGAGATATTTGTAACCTGCTGGGGCG 2097  
1981 GAACTCAAGATGATGACCCAGCTGGGAAGCCACAGAGATATTTGTAACCTGCTGGGGCG 2040  
2098 TGACACATGTGAGACCAATTTACTGATTTTGAATCTGCTGCTATGCTGATCTTCTC 2157  
2041 TGACACATGTGAGACCAATTTACTGATTTTGAATCTGCTGCTATGCTGATCTTCTC 2100  
2158 AACTATCTAAGATGAAAGAGAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAA 2217  
2101 AACTATCTAAGATGAAAGAGAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAA 2160  
2218 CACAAATTTAGTTTATCCCACTTCCAAATCACAATCCAAATTCAGATCCCTGGTTCA 2277  
2161 CACAAATTTAGTTTATCCCACTTCCAAATCACAATCCAAATTCAGATCCCTGGTTCA 2220  
2278 AGAGAAGTTCAATACACCCGAGTCTGGATCAAAATCTCAGGGCTTCATGGGAATTTCAAT 2337  
2221 AGAGAAGTTCAATACACCCGAGTCTGGATCAAAATCTCAGGGCTTCATGGGAATTTCAAT 2280  
2338 CACTCTGAAGATGAAATTTGAATATGAATAACAAAGGCTGGGAAGAGGAGGACTTG 2397

2281 CACTCTGAAGATGAAATTTGAATATGAATAACAAAGGCTGGAAGAGAGGAGACTTG 2340  
2398 AATGCTCTTACATTTGAAGATCTTCTTGTCTTGCATATCAAGTTGCCAAGGAATGGAA 2457  
2341 AATGCTCTTACATTTGAAGATCTTCTTGTCTTGCATATCAAGTTGCCAAGGAATGGAA 2400  
2458 TTTCTGGAATTTAAATGCTGTGTTCACAGACAGCTGGCGCCAGGAACTGTCTTGTTCACC 2517  
2401 TTTCTGGAATTTAAATGCTGTGTTCACAGACAGCTGGCGCCAGGAACTGTCTTGTTCACC 2460  
2518 CACGGGAAGTGTGAAGATATGTGAATTTGGATTTGGTTCGAGATATCATATGATGATTC 2577  
2461 CACGGGAAGTGTGAAGATATGTGAATTTGGATTTGGTTCGAGATATCATATGATGATTC 2520  
2578 AACTATGTTGTGAGGGCAATGCCGCTCTGCTGTAAATGATGAGGCCCCGAAAGCCTG 2637  
2521 AACTATGTTGTGAGGGCAATGCCGCTCTGCTGTAAATGATGAGGCCCCGAAAGCCTG 2580  
2638 TTTCAAGGCATCTACACCATTAAGAGTATCTCTGCTCATATGGAATATTTACTGTGGAA 2697  
2581 TTTGAAGGCATCTACACCATTAAGAGTATCTCTGCTCATATGGAATATTTACTGTGGAA 2640  
2698 ATCTTCTCACTTGTGTGAATCTTACCTGGCAATTCGGTTGATGCTAACTTCTACAAA 2757  
2641 ATCTTCTCACTTGTGTGAATCTTACCTGGCAATTCGGTTGATGCTAACTTCTACAAA 2700  
2758 CTGATTCAAAATGATTTAAATGATCAGCCATTTTATGCTACAGAAAGAAATATACAT 2817  
2701 CTGATTCAAAATGATTTAAATGATCAGCCATTTTATGCTACAGAAAGAAATATACAT 2760  
2818 ATAATGCAATCTGCTGGCTTTTGAATCAAGGAAACGGCCATCTTCCCTAAATTTGACT 2877  
2761 ATAATGCAATCTGCTGGCTTTTGAATCAAGGAAACGGCCATCTTCCCTAAATTTGACT 2820  
2878 TCGTTTTTAGATCTCAGCTGGCAGATGCAGAAAGCGATGTATCAGAACTGTGATGCG 2937  
2821 TCGTTTTTAGATCTCAGCTGGCAGATGCAGAAAGCGATGTATCAGAACTGTGATGCG 2880  
2938 CGTGTTCGGAATCTCTCACACCTTACAAACAGGCGACCTTTTCAGCAGAGAGATGAT 2997  
2881 CGTGTTCGGAATCTCTCACACCTTACAAACAGGCGACCTTTTCAGCAGAGAGATGAT 2940  
2998 TTGGGGCTACTCTCTCGCAGGCTCAGTCAAGATTCGTAGAGGAACTTTAGTTTATA 3057  
2941 TTGGGGCTACTCTCTCGCAGGCTCAGTCAAGATTCGTAGAGGAACTTTAGTTTATA 3000  
3058 AGGACTTCATCTCCCTCACTTCCCTAACAGGCTGTAGATTTACCAACACAGATTAATTT 3117  
3001 AGGACTTCATCTCCCTCACTTCCCTAACAGGCTGTAGATTTACCAACACAGATTAATTT 3060  
3118 CATCACTAAAAGAAAATCTATTATCACTGCTGCTTACACAGACTTTTCTCTAGAGCGG 3177  
3061 CATCACTAAAAGAAAATCTATTATCACTGCTGCTTACACAGACTTTTCTCTAGAGCGG 3120

## RESULT 4

US-09-919-408-1  
; Sequence 1, Application US/09919408  
; Patent No. US200207207A1  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Imclone Systems Incorporated  
; STREET: 180 Varick Street  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk



COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/919,408  
FILING DATE: 31-Jul-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/977,451  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US 07/906,397  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: US PCT/US92/05401  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: TW 81102961  
FILING DATE: 15-APR-1992  
APPLICATION NUMBER: US PCT/US92/02750  
FILING DATE: 02-APR-1992  
APPLICATION NUMBER: US 07/813,593  
FILING DATE: 24-DEC-1991  
APPLICATION NUMBER: US 07/793,065  
FILING DATE: 15-NOV-1991  
APPLICATION NUMBER: US 07/728,913  
FILING DATE: 28-JUN-1991  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Feit, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054

## INFORMATION FOR SEQ ID NO: 1:

## SEQUENCE CHARACTERISTICS:

LENGTH: 3453 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

FEATURE:

NAME/KEY: mat peptide

LOCATION: 112..3006

FEATURE:

NAME/KEY: sig\_peptide

LOCATION: 31..1111

FEATURE:

NAME/KEY: CDS

LOCATION: 31..3009

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-919-408-1

Query Match 63.6%; Score 2227.2; DB 9; Length 3453;

Best Local Similarity 79.6%; Pred. No. 0;

Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;

Qy	35	GGGACCCCGGCTCCGGAGCCATGCGCGTGGC---GCGGACCGGGGACCGTGC	91
Db	8	GGCTACCGCGGCTCCGGAGCCATGCGCGTGGCAGCGGACCGCGGCGTGC	67
Qy	92	CGTGTCTGTTTCTGCTGATGATTTGGACTATTACAAATCAAGTCTGCTG	151
Db	68	TGCTGTGTTTGTGCTGATGATTTGGAGCCGTTACAAACCAAGCCTGCTG	127
Qy	152	TGATCAAGTGTTTATCAATCATAGAACATGATTCACTAGTGGGAGTCAATCAT	211
Db	128	TGATCAAGTGTTTATCAAGTCAAGAACATGCTGATGCGGGAAGCCATCAT	187
Qy	212	CATATCCCATGGTATCAGAAATCCCGGGAAGACCTCGGCTGTGCGTTGAGACCCAGACT	271

Db	188	CGTACCGAATGGTGGAGGATCCCAGAAAGACCTCCAGTGTACCCCGAGGCGCCAGAGTG	247
Qy	272	CAGGACAGTGTACGAAGCTGCGCTGTGGAAGTGGATGATCTGCTCCATCACACTGC	331
Db	248	AAGGACGGTATATGAAGCGCCACCGTGGAGTGGCGAGTCTGGGTCCATCACCTGC	307
Qy	332	AAGTGTGTCGATGCGCCAGGGAACATTTCTGCTCTCGGTCTTTAAGCACAGCTCCC	391
Db	308	AAGTGCAGCTCGCCACCCAGGGGACCTTCTCCTCCTGGGTCTTTAAGCACAGCTCCC	367
Qy	392	TGAATTGCCAGCCACATTTTGAATTTACAAACAGAGGAGTGTTCATGCTCATTTGA	451
Db	368	TGGCTGCCAGCCGCACTTTGAATTTACAAACAGAGGAATCGTTTCCATGCCATCTGA	427
Qy	452	AAATGACAGAAACCCAAAGCTGGAGATACCTACTTTTATTCAGAGTGAAGCTACCAATT	511
Db	428	ACGTGACAGAGACCCAGGAGGAGATACCTACTCCATATTCAGAGGGAACGCGCAACT	487
Qy	512	ACACATATTTGTTTACAGTGAATTAAGAAATACCTGCTTTACACATTAAAGAGACCTT	571
Db	488	ACACAGTACTGTTTACAGTGAATTAAGAGATACACAGCTGTATGTCTAAGGAGACCTT	547
Qy	572	ACTTTAGAAAAATGAAAAACAGGACCCCTGGTCTGCTATCTGAGAGCGTTCAGAGC	631
Db	548	ACTTTAGAAAGTGAATAACAGGATGCACTGCTCTGCACTCCGAGGGTGTTCGGAGC	607
Qy	632	CGATCGTGAATGGTGTCTTTGCGATTACAGGGGGGAAAGCTGTAAAGAAAGATCCAG	691
Db	608	CCACTGTGGAGTGGTGTCTCTGCACTCCACAGGAAAGCTGTAAAGAAAGAGCCCTG	667
Qy	692	CTGTTGTTAAAGGAGGAAAAAGTGTCTTATGAAATTTTGGGACGGACATAGGTGCT	751
Db	668	CTGTTGTGAGAAAGGAGGAAAAAGGTACTTTCATGAGTTGTTTCGGAACAGACATCAGATGCT	727
Qy	752	GTGCCAGAAATGAATCGGCGAGGAATGCACCGGCTGTTCACAAATAGATCTAATAACAA	811
Db	728	GTGCTAGAAATGCACTGGCGCGGAATGCACAGCTGTTCCACATAGATCTAATAACAGG	787
Qy	812	CTCCTCAGACACATTCGCCAAATTTATTTCTTAAAGTAGGGGAAACCTTTATGGATAAGT	871
Db	788	CTCCTCAGAGACACTGCCCCAGTTTATTCCTGAAAGTGGGGGAAACCTTTGTGGATCAGT	847
Qy	872	GCAAAGCTGTTTCATGTGAACCATCGATTGCGGCTCACTCGGGAATTAGAAAAACAAAGCAC	931
Db	848	GTAAGGCCATCCATGTGAACCATCGATTGCGGCTCACTCGGAGCTGGGAACAAAAGCCC	907
Qy	932	TCGAGGAGGCAACTACTTTGAGATGAGTACTTATTCAAACAAACAGACTATGATACGGA	991
Db	908	TGGAGGAGGCGAGCTACTTTGAGATGAGTACTTACTCCAAACAGGAGCCATGATTCGGA	967
Qy	992	TTCTGTTTGTCTTTGTATCATCAGTGGCAAGAAACGACACCGGATATACACTTGTCTCT	1051
Db	968	TTCTTTGGCTTTGTGTCTTCGTGGGAGGAACGACACCGGATATTACACCTGCTCTT	1027
Qy	1052	CTTCAAGAGCATCCAGTCAATCAGCTTTGGTTTCACTGTTAGGAAGGATTTTATAATG	1111
Db	1028	CCTCAAGAGCAACCCAGCCAGTCCAGCTTTGGTGGACCATCTAGAAAAAGGGTTATAAAG	1087
Qy	1112	CTACCAATTCAGTGAAGATTTGAAATTCACCAATATGAGAGTGTGTTTCTGTCGA	1171
Db	1088	CTACCACTCGCAAGAGAGATGAAATTCGCGGTACGAAAGTTCTGCTTCTCAGTCA	1147
Qy	1172	GGTTTAAAGCCTACCCCAATCAGATGTACGTGGACCTCTCTCGAAAAATCATTTCTCTT	1231
Db	1148	GGTTTAAAGCGTACCCAGAAATCCGATGCACTGGATCTCTCTCAAGCCTCATTTCTCTT	1207
Qy	1232	GTGAGCAAAAGGGTCTTTGATAACCGGATACAGCATATCCAAAGTTTTCGAATCATAGCAC	1291
Db	1208	GTGAACAGAGAGGCTTGAGGATGGGTACAGCATATCTAAATTTTTCGATCATAGAACA	1267
Qy	1292	AGCCAGGAGATATATATTCATCCATGAGAAATGATGATGATGATGATGATGATGATGATGAT	1351
Db	1268	AGCCAGGAGATATATATTCATCCATGAGAAATGATGATGATGATGATGATGATGATGATGAT	1327



1352 CGCTGAATATAAGAGGAAACCTCAAGTGTCTCGAGAAGCATCGGCAAGTCAGCGCTCCT 1411  
1328 CGCTGAATATAAGAGGAAACCTCAAGTGTCTAGCAATGCTCAGCCAGCCAGCGCTCCT 1387  
1412 GTTCTCGGATGATACCAATACCAATCTTGGACCTGGAGAGAGTGTTCAGACAAGTCTC 1471  
1388 GTTCTCTGATGGCTACCGCTACCTCTTGGACCTGGAGAGAGTGTTCGGAACAATCTC 1447  
1472 CCAACTGCCACAGAAGAGATACACAGAAGGAGTCTGGGAATAGAAAGGCTTACAGAAAAGTGT 1531  
1448 CCAATTTGCAGCGAGGAATCCCAAGAGAGTGTGGGAATAAAAGGCTTAAAGAAAAGTGT 1507  
1532 TTGGACAGTGGGTGTCTGAGCAGTACTCTAAACATAGAGTGAAGCATAAAAGGCTTCTGG 1591  
1508 TTGGCCAGTGGGTGTCTGAGCAGTACTCTAAATATAGAGTGAAGCGCGGGAAGGCTTCTGG 1567  
1592 TCAAGTGTCTGATACAAATTCCTTTGGGACATCTTGTGAGAGAGTCTTTTAAATCTCTC 1651  
1568 TCAAAATGTCTGTGCTACAAATTTCTATGGGACGCTCTTGGGAACCAATCTTTTAAATCTC 1627  
1652 CAGGCCCTTCCCTTTTATCCCAAGACAACATCTTCTATGGGACGCTCTTGGGAACCAATCTTTTAAATCTC 1711  
1628 CAGGCCCTTCCCTTTTATCCCAAGACAACATCTTCTATGGGACGCTCTTGGGCTCTGTC 1687  
1712 TCCTCTTCAATGTCTGTTTTAAACCTCTCAATTTGTCAAGTACAAAAGCAATTTAGT 1771  
1688 TCCCTCTTCAATGTCTGTTTTCAATTTGTCTGATCTGCCACAATATCAAAAAGCAATTTAGT 1747  
1772 ATGAAGCCAGCTACAGATGGTACAGGTGACCGGCTCTCAGATATAGTACTTCTACG 1831  
1748 ACAGAGTCAAGTGTGATGATCTCAATGGAGTTCGAGAGAGAGTCTTCTACG 1807  
1832 TTGATTTCAAGAAATATGATATGATCTCAATGGAGTTCGAGAGAGAGTCTTCTACG 1891  
1808 TTGATTTCAAGAGTATGATATGATCTCAATGGAGTTCGAGAGAGAGTCTTCTACG 1867  
1892 TTGGAAGTACTAGATCAGGTGCTTTTGGAAAAGTGTGATGACGCAACAGCTTATGGAA 1951  
1868 TTGGAAGTCTGGGCTCTGGGCTTTTGGGAGGTGATGAAAGCGCAACGCTATGGCA 1927  
1952 TTAGCAAAACAGGAGTCTCAATCCAGTGTGCGGTCAAAATGTGAAAGAAAAGCAGACA 2011  
1928 TTAGTAAACAGGAGTCTCAATTTAGTGGCGGTGAGATGTGAAAGAGAAAGCTGACA 1987  
2012 GCTCTGAAGAGAGGAGTCTATGTGAGAACTCAAGATGATGACCCAGTGGGAAGCCAG 2071  
1988 GCTGTGAAAAAGAGTCTCTATGTGCGAGCTCAAAATGATGACCCACCTGGGACACCATG 2047  
2072 AGAATATTTGAACTGTGGGGGTGACACTGTGAGGAGTCTTACTTGTGATTTTGTG 2131  
2048 ACAATCTGTGAATCTGTGGGGGTGACACTGTGAGGAGTCTTACTTGTGATTTTGTG 2107  
2132 AATATCTGTGTGATCTTCTCAACTATCTAAGAAAGTAAAGAGAAAATTTTCA 2191  
2108 AATATTTGTGTAAGTGTGACCTCTCACTACTTAAGAAAGTAAAGAGAAAGTTCACA 2167  
2192 GGACTGTGACAGAGATTTTCAAGGAACAAATTTTACGCTTTTACCCACTTTTCAATCAC 2251  
2168 GGACATGGACAGAGATTTTAAAGGAACATAATTTTCACTTTTACCCACTTTTCCAGGCAC 2227  
2252 ATCCAAATTCAGCATGCTGTGTTCAAGAGAGTTCAGATACACCCGAGTCTGGATCAAA 2311  
2228 ATTCAAAATTCAGCATGCTGTGTTCAAGAGAGTTCAGTATACCCGCGCTTGGATCAGC 2287  
2312 TCTCAGGCTTCTATGGGAATTTCAATTTCTCTGAGAGTGAATTTGAATATGAAGAACCAA 2371  
2288 TCTCAGGCTTCAAAGGGAATTTCAATTTCTGAGAGTGAATTTGAATATGAAGAACCAA 2347  
2372 AAAGGCTG-----GAAGAGAGGAGGACTTGAATGTGCTTACATTTGAAGATCTTCTTT 2425  
2348 AGAGGCTGGCAGAGAGAGAGGAGGAGATTTGAAAGTGTCTGACGCTTGTGAAGCTCTTTT 2407

2426 GCTTTGCTATCAAGTGTGCCCCAAGGAATGGAATTTCTGGAAATTTAAGTCGTGTCTTCA 2485  
2408 GCTTTGCTATCAAGTGTGCCCCAAGGAATGGAATTTCTGGAAATTTAAGTCGTGTCTTCA 2467  
2486 GAGACCTGGCCGCGCAGGAAACGTCTTTGTTCACCCACGCGGAAAGTGTGGAATATGTGACT 2545  
2468 GAGACCTGGCAGCAGGAAATGTGTGGTTCACCCACGCGGAAAGTGTGGAATATGTGACT 2527  
2546 TTGGATTTGGTCTGAGATATCATGAGTGTTCACAACTATGTTCAGGGGCAATGCCGCTC 2605  
2528 TTGGACTGGCCGCGCAGACATCTCTGAGCGACTCTCAGCTACGTCTGTCAGGGGCAACGCGGC 2587  
2606 TGCTCTGAAAAATGAGTGGCCCGGAAAGCCTGTGTTGAAGCATCTACACCATTAAGAGTGT 2665  
2588 TGCCGCTGAGTGTGAGTGGCACCCGAGAGCTTATTTGAAGGATCTACACAAATCAAGAGTGT 2647  
2666 ATGTCTGGTCTATATGGAATATTTACTGTGGGAAATCTTCTCAGTTGGTGTGAATCCTTACC 2725  
2648 AGTCTGGTCTTACGCGCATCTCTCTGGGAGATATTTTCACTGGGTGTGAAACCTTACC 2707  
2726 CTGGCATTTCCGTTGTATGCTTCTTAACAATGATTTCAAAATGGAATTTTAAATGGAATC 2785  
2708 CTGGCATTTCTGTGCGAGCTTCTTCTATAAACTGATTCAGAGTGGATTTTAAATGGAATC 2767  
2786 AGCCATTTTATGCTACAGAGAAATATACATTAATATGCAATCTCTGCTGGGCTTTTACT 2845  
2768 AGCCATTTTATGCTACAGAGAAATATGATTTGTAATGCAATCTCTGCTGGGCTTTTACT 2827  
2846 CAAGAAAACGCGCATCTTCCCTTAAATTTGACTTTCGTTTTTAGGATGTGAGTCTGCGAGATG 2905  
2828 CAAGNAAGGCGCATCTTCCCAACCTGACTTCTATTTTAGGATGTGAGTCTGCGAGAG 2887  
2906 CAGAAAGCAGATGATCAGAAATGTGATGGCGGTGTTTCGGAATGTCTCTCACACCTACC 2965  
2888 CAGAAAGAG-----CATGATACAGAAATCATCCATCATCTACC 2924  
2966 AAAACAGGCGACTTTACAGAGAGATGGAATTTGGGGTACTCTCTCCGAGGCTCAGG 3025  
2925 AAAACAGGCGGCTCTCAGCAGAGAGGCGGCTCAGAGC-CCAGTCCGCAAGCGGCGAG 2983  
3026 TCGAAGATTCGTAGAGGAAATTTAGTTTTTAAGGACTTCTCCCTTCCACCTATCTCTTAA 3085  
2984 T-GAAGATTCAGAGAAAGAAATTTAGCGAGGCGCTTGGAGCGCTTGGACCCCGC-----CACCTAG 3038  
3086 CAGGCTGTAGATTCACAAAACAAAGTAATTTTATCACTTAAAGAAAATCTATTTATCAAC 3145  
3039 CAGGCTGTAGAGCGCAGAGCGCAAGATTTAGCCTCGGCTCT--GAGGAAGCGCCCTCAGCG 3096  
3146 TGCTGCTTCAACAGACTTTTCTAGAAAGCGCTCTGCGTTTACTCTTGTGTTTCAAAAGGGA 3205  
3097 CGTTGCTTCCGCTGGACTTTTCTAGATGCTGTCTGCCATTTACTC-----CAAAGTGA 3149  
3206 CTTTGTAAATCAATCATCTCTGTCACAGCAGAGAGCTGATATGAACTTTTATTTG 3265  
3150 CTTTGTAAATCAAACTCTCTCGCAGCGGAGAGCAATAATGAGACTTTGTTG 3209  
3266 GAGCATTTGATCTGCAATCAAGCGCTTCTCAGCGCGCTTGTGAGTGAATTTGACTTTATTTG 3325  
3210 TGAGCCCGCTACCTGGGGCTTTTCCAGAGCTTGGGGGAAAGCCATGTATCTGAAA 3269  
3326 TACAGTATTTTCTGTAATAATCAAAAACAA-----AGCATTTTGTCTAAGGAGAAGC 3378  
3270 TATAGTATTTCTGTAATACGTGAACAAACAAACCCCGTTTTTGTCTAAGGGAAGC 3329  
3379 TAAATGATTTTTT--AAGTCTATGTTTTTAAATAATATGTAATTTTTCAGTATTTAG 3436  
3330 TAAATGATTTTTTAAATAATCTATGTTTTTAAATAATGTAATTTTCTCTATTTAG 3389  
3437 TGATATTTTTTATGGGTGGGAATAAATTTCTCTACAGAAAGAAAATAAATAAATAAATAA 3496  
3390 TGATATTTTTTATGATGGAATTAACCTTTCTACTGTAAAAATAAATAAATAAATAAATAA 3449  
3497 AAAA 3500

```
Db      3450 AAAA 3453
|||||
RESULT 5
US-09-872-136-1
; Sequence 1, Application US/09872136
; Patent No. US20020119545A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESS: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/872,136
; FILING DATE: 01-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/208,786
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/09/021,324
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/07/977,451
; FILING DATE: 1992-11-19
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-7P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3453 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; NAME/KEY: mat_peptide
; LOCATION: 112..3006
; FEATURE:
;
; NAME/KEY: sig_peptide
; LOCATION: 31..111
; FEATURE:
;
; NAME/KEY: CDS
; LOCATION: 31..3009
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-872-136-1
Query Match      63.6%; Score 2227.2; DB 10; Length 3453;
Best Local Similarity 79.6%; Pred. No. 0;
Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;
QY 35 GGGGACCCCGGCTCCGAGAGCCATGCCGGGTGGC---GGCGACGGGGCCACCGTGC 91
Db 8 GGCTACCGCGGCTCCGAGAGCCATGCCGGGTGGCAGCGAGCCAGCCGCGGCTGC 67
QY 92 CGCTGCTGTTGTTTCTGCAATGATATTGGGACTATTACAAATCAAGATCTGCCTG 151
Db 68 TGCTGCTGTTGTTTGTCAATGATCTTGAGACCGTTACAAACCAAGACCTGCTG 127
QY 152 TGATCAAGTGTGTTTAAATCAATCATAGAAACAATGATTCATCTAGTGGGAAAGTCA 211
Db 128 TGATCAAGTGTGTTTAAATCAATCATAGAAACAATGATTCATCTAGTGGGAAAGTCA 187
QY 212 CATATCCATGCTATCAGATCCCGGAGAGCTCCGGGTGCGTGTGAGACCCAGAGCT 271
Db 188 CGTACCGAATGCTGCGAGGATCCCGAGAGACCTCCAGTGTACCCGAGGCGCCAGAG 247
QY 272 CAGGACAGTGTACGAGCTGCGGTGGAAGTGTATCTCTCTTCCATCACAATGC 331
Db 248 AAGGACGCTATATGAGCGGCCACCGTGGAGGTGGCCGAGTCTGGTCCATCCCTGC 307
QY 332 AAGTGTGCTGATGCCCGGAGGAAATTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 391
Db 308 AAGTGTGCTGATGCCCGGAGGAAATTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 367
QY 392 TGAATGCCAGCCACATTTTGAATTTACAAACAGAGGAGTGTGTTCCATGTCATTTGA 451
Db 368 TGGGCTGCCAGCCGACCTTTGATTTTACAAACAGAGGAAATCGTTTCCATGCGCATTTGA 427
QY 452 AAATCAGACAGAACCCAGCTGAGAGTACCTACTTTTATTCAGAGTGAAGTCACTCAAT 511
Db 428 ACGTGACAGAGACCCAGGACGAGAGATACCTACTTCCATATTCAGAGCGGAAACCGG 487
QY 512 ACACAAATTTGTACAGTGTATAGAAATACCTCTGCTTTACACATTTAAGAGACCTT 571
Db 488 ACAGTACTGTTACAGTGTATGTAAGAGATACACAGCTGTATGTGTAGAGAGCTT 547
QY 572 ACTTTAGAAAAATGGAAAAACAGGACGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 631
Db 548 ACTTTAGGAAGATGGAAAAACAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 607
QY 632 CGATCGTGAATGGTGTCTTTCGATTTACAGGGGGAAGCTGTAAAGAGAAAGTCCAG 691
Db 608 CCACTGTGGAGTGGTGTCTCTGAGCTCCCAAGGAAAGCTGTAAAGAGAAAGGCTG 667
QY 692 CTGTTGTTAAAAAGGAGAAAAAGTCTTCAATGATTTTGGGACCGACATAAGTGTCT 751
Db 668 CTGTTGTCAGAAAGGAGAAAGTACTTCAATGATTTTGGGACCGACATAAGTGTCT 727
QY 752 GTGCCAGAAATGAACCTGGGAGGGAATGACAGGCTGTTCACAAATAGATCTAAATCAA 811
Db 728 GTGCTAGAAATGCACTGGGCGCGAATGCAACCAAGCTGTTCACCAATAGATCTAAAC 787
QY 812 CTCCTCAGACCAATGTCACAAATTTTCTTAAAGTAGGGGAACCTTATGGATAAGT 871
Db 788 CTCCTCAGAGCACCTGCCCGGAGTTTATTCCTGAAAGTGGGGAAACCTTGTGGATCAG 847
QY 872 GCAAGCTGTTTCATGTGAACCATGATTCGGGCTCACCTCGGAAATTTAGAAAAACAAG 931
Db 848 GTAAGGCCATCCATGTGAACCATGATTCGGGCTCACCTCGGAGCTGGGAAGACAAGCC 907
QY 932 TCGAGGAGGCAACTACTTTTGAGATGAGTACCTATTCAACAAACAGAACTATATACGA 991
```

1008 TGAGAGGAGGAGCTACTTTGAGATGAGTACTTACCAAAACAGGACCATGATTCGGA 967  
1009  
1010  
1011  
1012  
1013  
1014  
1015  
1016  
1017  
1018  
1019  
1020  
1021  
1022  
1023  
1024  
1025  
1026  
1027  
1028  
1029  
1030  
1031  
1032  
1033  
1034  
1035  
1036  
1037  
1038  
1039  
1040  
1041  
1042  
1043  
1044  
1045  
1046  
1047  
1048  
1049  
1050  
1051  
1052  
1053  
1054  
1055  
1056  
1057  
1058  
1059  
1060  
1061  
1062  
1063  
1064  
1065  
1066  
1067  
1068  
1069  
1070  
1071  
1072  
1073  
1074  
1075  
1076  
1077  
1078  
1079  
1080  
1081  
1082  
1083  
1084  
1085  
1086  
1087  
1088  
1089  
1090  
1091  
1092  
1093  
1094  
1095  
1096  
1097  
1098  
1099  
1100  
1101  
1102  
1103  
1104  
1105  
1106  
1107  
1108  
1109  
1110  
1111  
1112  
1113  
1114  
1115  
1116  
1117  
1118  
1119  
1120  
1121  
1122  
1123  
1124  
1125  
1126  
1127  
1128  
1129  
1130  
1131  
1132  
1133  
1134  
1135  
1136  
1137  
1138  
1139  
1140  
1141  
1142  
1143  
1144  
1145  
1146  
1147  
1148  
1149  
1150  
1151  
1152  
1153  
1154  
1155  
1156  
1157  
1158  
1159  
1160  
1161  
1162  
1163  
1164  
1165  
1166  
1167  
1168  
1169  
1170  
1171  
1172  
1173  
1174  
1175  
1176  
1177  
1178  
1179  
1180  
1181  
1182  
1183  
1184  
1185  
1186  
1187  
1188  
1189  
1190  
1191  
1192  
1193  
1194  
1195  
1196  
1197  
1198  
1199  
1200  
1201  
1202  
1203  
1204  
1205  
1206  
1207  
1208  
1209  
1210  
1211  
1212  
1213  
1214  
1215  
1216  
1217  
1218  
1219  
1220  
1221  
1222  
1223  
1224  
1225  
1226  
1227  
1228  
1229  
1230  
1231  
1232  
1233  
1234  
1235  
1236  
1237  
1238  
1239  
1240  
1241  
1242  
1243  
1244  
1245  
1246  
1247  
1248  
1249  
1250  
1251  
1252  
1253  
1254  
1255  
1256  
1257  
1258  
1259  
1260  
1261  
1262  
1263  
1264  
1265  
1266  
1267  
1268  
1269  
1270  
1271  
1272  
1273  
1274  
1275  
1276  
1277  
1278  
1279  
1280  
1281  
1282  
1283  
1284  
1285  
1286  
1287  
1288  
1289  
1290  
1291  
1292  
1293  
1294  
1295  
1296  
1297  
1298  
1299  
1300  
1301  
1302  
1303  
1304  
1305  
1306  
1307  
1308  
1309  
1310  
1311  
1312  
1313  
1314  
1315  
1316  
1317  
1318  
1319  
1320  
1321  
1322  
1323  
1324  
1325  
1326  
1327  
1328  
1329  
1330  
1331  
1332  
1333  
1334  
1335  
1336  
1337  
1338  
1339  
1340  
1341  
1342  
1343  
1344  
1345  
1346  
1347  
1348  
1349  
1350  
1351  
1352  
1353  
1354  
1355  
1356  
1357  
1358  
1359  
1360  
1361  
1362  
1363  
1364  
1365  
1366  
1367  
1368  
1369  
1370  
1371  
1372  
1373  
1374  
1375  
1376  
1377  
1378  
1379  
1380  
1381  
1382  
1383  
1384  
1385  
1386  
1387  
1388  
1389  
1390  
1391  
1392  
1393  
1394  
1395  
1396  
1397  
1398  
1399  
1400  
1401  
1402  
1403  
1404  
1405  
1406  
1407  
1408  
1409  
1410  
1411  
1412  
1413  
1414  
1415  
1416  
1417  
1418  
1419  
1420  
1421  
1422  
1423  
1424  
1425  
1426  
1427  
1428  
1429  
1430  
1431  
1432  
1433  
1434  
1435  
1436  
1437  
1438  
1439  
1440  
1441  
1442  
1443  
1444  
1445  
1446  
1447  
1448  
1449  
1450  
1451  
1452  
1453  
1454  
1455  
1456  
1457  
1458  
1459  
1460  
1461  
1462  
1463  
1464  
1465  
1466  
1467  
1468  
1469  
1470  
1471  
1472  
1473  
1474  
1475  
1476  
1477  
1478  
1479  
1480  
1481  
1482  
1483  
1484  
1485  
1486  
1487  
1488  
1489  
1490  
1491  
1492  
1493  
1494  
1495  
1496  
1497  
1498  
1499  
1500  
1501  
1502  
1503  
1504  
1505  
1506  
1507  
1508  
1509  
1510  
1511  
1512  
1513  
1514  
1515  
1516  
1517  
1518  
1519  
1520  
1521  
1522  
1523  
1524  
1525  
1526  
1527  
1528  
1529  
1530  
1531  
1532  
1533  
1534  
1535  
1536  
1537  
1538  
1539  
1540  
1541  
1542  
1543  
1544  
1545  
1546  
1547  
1548  
1549  
1550  
1551  
1552  
1553  
1554  
1555  
1556  
1557  
1558  
1559  
1560  
1561  
1562  
1563  
1564  
1565  
1566  
1567  
1568  
1569  
1570  
1571  
1572  
1573  
1574  
1575  
1576  
1577  
1578  
1579  
1580  
1581  
1582  
1583  
1584  
1585  
1586  
1587  
1588  
1589  
1590  
1591  
1592  
1593  
1594  
1595  
1596  
1597  
1598  
1599  
1600  
1601  
1602  
1603  
1604  
1605  
1606  
1607  
1608  
1609  
1610  
1611  
1612  
1613  
1614  
1615  
1616  
1617  
1618  
1619  
1620  
1621  
1622  
1623  
1624  
1625  
1626  
1627  
1628  
1629  
1630  
1631  
1632  
1633  
1634  
1635  
1636  
1637  
1638  
1639  
1640  
1641  
1642  
1643  
1644  
1645  
1646  
1647  
1648  
1649  
1650  
1651  
1652  
1653  
1654  
1655  
1656  
1657  
1658  
1659  
1660  
1661  
1662  
1663  
1664  
1665  
1666  
1667  
1668  
1669  
1670  
1671  
1672  
1673  
1674  
1675  
1676  
1677  
1678  
1679  
1680  
1681  
1682  
1683  
1684  
1685  
1686  
1687  
1688  
1689  
1690  
1691  
1692  
1693  
1694  
1695  
1696  
1697  
1698  
1699  
1700  
1701  
1702  
1703  
1704  
1705  
1706  
1707  
1708  
1709  
1710  
1711  
1712  
1713  
1714  
1715  
1716  
1717  
1718  
1719  
1720  
1721  
1722  
1723  
1724  
1725  
1726  
1727  
1728  
1729  
1730  
1731  
1732  
1733  
1734  
1735  
1736  
1737  
1738  
1739  
1740  
1741  
1742  
1743  
1744  
1745  
1746  
1747  
1748  
1749  
1750  
1751  
1752  
1753  
1754  
1755  
1756  
1757  
1758  
1759  
1760  
1761  
1762  
1763  
1764  
1765  
1766  
1767  
1768  
1769  
1770  
1771  
1772  
1773  
1774  
1775  
1776  
1777  
1778  
1779  
1780  
1781  
1782  
1783  
1784  
1785  
1786  
1787  
1788  
1789  
1790  
1791  
1792  
1793  
1794  
1795  
1796  
1797  
1798  
1799  
1800  
1801  
1802  
1803  
1804  
1805  
1806  
1807  
1808  
1809  
1810  
1811  
1812  
1813  
1814  
1815  
1816  
1817  
1818  
1819  
1820  
1821  
1822  
1823  
1824  
1825  
1826  
1827  
1828  
1829  
1830  
1831  
1832  
1833  
1834  
1835  
1836  
1837  
1838  
1839  
1840  
1841  
1842  
1843  
1844  
1845  
1846  
1847  
1848  
1849  
1850  
1851  
1852  
1853  
1854  
1855  
1856  
1857  
1858  
1859  
1860  
1861  
1862  
1863  
1864  
1865  
1866  
1867  
1868  
1869  
1870  
1871  
1872  
1873  
1874  
1875  
1876  
1877  
1878  
1879  
1880  
1881  
1882  
1883  
1884  
1885  
1886  
1887  
1888  
1889  
1890  
1891  
1892  
1893  
1894  
1895  
1896  
1897  
1898  
1899  
1900  
1901  
1902  
1903  
1904  
1905  
1906  
1907  
1908  
1909  
1910  
1911  
1912  
1913  
1914  
1915  
1916  
1917  
1918  
1919  
1920  
1921  
1922  
1923  
1924  
1925  
1926  
1927  
1928  
1929  
1930  
1931  
1932  
1933  
1934  
1935  
1936  
1937  
1938  
1939  
1940  
1941  
1942  
1943  
1944  
1945  
1946  
1947  
1948  
1949  
1950  
1951  
1952  
1953  
1954  
1955  
1956  
1957  
1958  
1959  
1960  
1961  
1962  
1963  
1964  
1965  
1966  
1967  
1968  
1969  
1970  
1971  
1972  
1973  
1974  
1975  
1976  
1977  
1978  
1979  
1980  
1981  
1982  
1983  
1984  
1985  
1986  
1987  
1988  
1989  
1990  
1991  
1992  
1993  
1994  
1995  
1996  
1997  
1998  
1999  
2000  
2001  
2002  
2003  
2004  
2005  
2006  
2007  
2008  
2009  
2010  
2011  
2012  
2013  
2014  
2015  
2016  
2017  
2018  
2019  
2020  
2021  
2022  
2023  
2024  
2025  
2026  
2027  
2028  
2029  
2030  
2031  
2032  
2033  
2034  
2035  
2036  
2037  
2038  
2039  
2040  
2041  
2042  
2043  
2044  
2045  
2046  
2047  
2048  
2049  
2050  
2051  
2052  
2053  
2054  
2055  
2056  
2057  
2058  
2059  
2060  
2061  
2062  
2063  
2064  
2065  
2066  
2067  
2068  
2069  
2070  
2071  
2072  
2073  
2074  
2075  
2076  
2077  
2078  
2079  
2080  
2081  
2082  
2083  
2084  
2085  
2086  
2087  
2088  
2089  
2090  
2091  
2092  
2093  
2094  
2095  
2096  
2097  
2098  
2099  
2100  
2101  
2102  
2103  
2104  
2105  
2106  
2107  
2108  
2109  
2110  
2111  
2112  
2113  
2114  
2115  
2116  
2117  
2118  
2119  
2120  
2121  
2122  
2123  
2124  
2125  
2126  
2127  
2128  
2129  
2130  
2131  
2132  
2133  
2134  
2135  
2136  
2137  
2138  
2139  
2140  
2141  
2142  
2143  
2144  
2145  
2146  
2147  
2148  
2149  
2150  
2151  
2152  
2153  
2154  
2155  
2156  
2157  
2158  
2159  
2160  
2161  
2162  
2163  
2164  
2165  
2166  
2167  
2168  
2169  
2170  
2171  
2172  
2173  
2174  
2175  
2176  
2177  
2178  
2179  
2180  
2181  
2182  
2183  
2184  
2185  
2186  
2187  
2188  
2189  
2190  
2191  
2192  
2193  
2194  
2195  
2196  
2197  
2198  
2199  
2200  
2201  
2202  
2203  
2204  
2205  
2206  
2207  
2208  
2209  
2210  
2211  
2212  
2213  
2214  
2215  
2216  
2217  
2218  
2219  
2220  
2221  
2222  
2223  
2224  
2225  
2226  
2227  
2228  
2229  
2230  
2231  
2232  
2233  
2234  
2235  
2236  
2237  
2238  
2239  
2240  
2241  
2242  
2243  
2244  
2245  
2246  
2247  
2248  
2249  
2250  
2251  
2252  
2253  
2254  
2255  
2256  
2257  
2258  
2259  
2260  
2261  
2262  
2263  
2264  
2265  
2266  
2267  
2268  
2269  
2270  
2271  
2272  
2273  
2274  
2275  
2276  
2277  
2278  
2279  
2280  
2281  
2282  
2283  
2284  
2285  
2286  
2287  
2288  
2289  
2290  
2291  
2292  
2293  
2294  
2295  
2296  
2297  
2298  
2299  
2300  
2301  
2302  
2303  
2304  
2305  
2306  
2307  
2308  
2309  
2310  
2311  
2312  
2313  
2314  
2315  
2316  
2317  
2318  
2319  
2320  
2321  
2322  
2323  
2324  
2325  
2326  
2327  
2328  
2329  
2330  
2331  
2332  
2333  
2334  
2335  
2336  
2337  
2338  
2339  
2340  
2341  
2342  
2343  
2344  
2345  
2346  
2347  
2348  
2349  
2350  
2351  
2352  
2353  
2354  
2355  
2356  
2357  
2358  
2359  
2360  
2361  
2362  
2363  
2364  
2365  
2366  
2367  
2368  
2369  
2370  
2371  
2372  
2373  
2374  
2375  
2376  
2377  
2378  
2379  
2380  
2381  
2382  
2383  
2384  
2385  
2386  
2387  
2388  
2389  
2390  
2391  
2392  
2393  
2394  
2395  
2396  
2397  
2398  
2399  
2400  
2401  
2402  
2403  
2404  
2405  
2406  
2407  
2408  
2409  
2410  
2411  
2412  
2413  
2414  
2415  
2416  
2417  
2418  
2419  
2420  
2421  
2422  
2423  
2424  
2425  
2426  
2427  
2428  
2429  
2430  
2431  
2432  
2433  
2434  
2435  
2436  
2437  
2438  
2439  
2440  
2441  
2442  
2443  
2444  
2445  
2446  
2447  
2448  
2449  
2450  
2451  
2452  
2453  
2454  
2455  
2456  
2457  
2458  
2459  
2460  
2461  
2462  
2463  
2464  
2465  
2466  
2467  
2468  
2469  
2470  
2471  
2472  
2473  
2474  
2475  
2476  
2477  
2478  
2479  
2480  
2481  
2482  
2483  
2484  
2485  
2486  
2487  
2488  
2489  
2490  
2491  
2492  
2493  
2494  
2495  
2496  
2497  
2498  
2499  
2500  
2501  
2502  
2503  
2504  
2505  
2506  
2507  
2508  
2509  
2510  
2511  
2512  
2513  
2514  
2515  
2516  
2517  
2518  
2519  
2520  
2521  
2522  
2523  
2524  
2525  
2526  
2527  
2528  
2529  
2530  
2531  
2532  
2533  
2534  
2535  
2536  
2537  
2538  
2539  
2540  
2541  
2542  
2543  
2544  
2545  
2546  
2547  
2548  
2549  
2550  
2551  
2552  
2553  
2554  
2555  
2556  
2557  
2558  
2559  
2560  
2561  
2562  
2563  
2564  
2565  
2566  
2567  
2568  
2569  
2570  
2571  
2572  
2573  
2574  
2575  
2576  
2577  
2578  
2579  
2580  
2581  
2582  
2583  
2584  
2585  
2586  
2587  
2588  
2589  
2590  
2591  
2592  
2593  
2594  
2595  
2596  
2597  
2598  
2599  
2600  
2601  
2602  
2603  
2604  
2605  
2606  
2607  
2608  
2609  
2610  
2611  
2612  
2613  
2614  
2615  
2616  
2617  
2618  
2619  
2620  
2621  
2622  
2623  
2624  
2625  
2626  
2627  
2628  
2629  
2630  
2631  
2632  
2633  
2634  
2635  
2636  
2637  
2638  
2639  
2640  
2641  
2642  
2643  
2644  
2645  
2646  
2647  
2648  
2649  
2650  
2651  
2652  
2653  
2654  
2655  
2656  
2657  
2658  
2659  
2660  
2661  
2662  
2663  
2664  
2665  
2666  
2667  
2668  
2669  
2670  
2671  
2672  
2673  
2674  
2675  
2676  
2677  
2678  
2679  
2680  
2681  
2682  
2683  
2684  
2685  
2686  
2687  
2688  
2689  
2690  
2691  
2692  
2693  
2694  
2695  
2696  
2697  
2698  
2699  
2700  
2701  
2702  
2703  
2704  
2705  
2706  
2707  
2708  
2709  
2710  
2711  
2712  
2713  
2714  
2715  
2716  
2717  
2718  
2719  
2720  
2721  
2722  
2723  
2724  
2725  
2726  
2727  
2728  
2729  
2730  
2731  
2732  
2733  
2734  
2735  
2736  
2737  
2738  
2739  
2740  
2741  
2742  
2743  
2744  
2745  
2746  
2747  
2748  
2749  
2750  
2751  
2752  
2753  
2754  
2755  
2756  
2757  
2758  
2759  
2760  
2761  
2762  
2763  
2764  
2765  
2766  
2767  
2768  
2769  
2770  
2771  
2772  
2773  
2774  
2775  
2776  
2777  
2778  
2779  
2780  
2781  
2782  
2783  
2784  
2785  
2786  
2787  
2788  
2789  
2790  
2791  
2792  
2793  
2794  
2795  
2796  
2797  
2798  
2799  
2800  
2801  
2802  
2803  
2804  
2805  
2806  
2807  
2808  
2809  
2810  
2811  
2812  
2813  
2814  
2815  
2816  
2817  
2818  
2819  
2820  
2821  
2822  
2823  
2824  
2825  
2826  
2827  
2828  
2829  
2830  
2831  
2832  
2833  
2834  
2835  
2836  
2837  
2838  
2839  
2840  
2841  
2842  
2843  
2844  
2845  
2846  
2847  
2848  
2849  
2850  
2851  
2852  
2853  
2854  
2855  
2856  
2857  
2858  
2859  
2860  
2861  
2862  
2863  
2864  
2865  
2866  
2867  
2868  
2869  
2870  
2871  
2872  
2873  
2874  
2875  
2876  
2877  
2878  
2879  
2880  
2881  
2882  
2883  
2884  
2885  
2886  
2887  
2888  
2889  
2890  
2891  
2892  
2893  
2894  
2895  
2896  
2897  
2898  
2899  
2900  
2901  
2902  
2903  
2904  
2905  
2906  
2907  
2908  
2909  
2910  
2911  
2912  
2913  
2914  
2915  
2916  
2917  
2918  
2919  
2920  
2921  
2922  
2923  
2924  
2925  
2926  
2927  
2928  
2929  
2930  
2931  
2932  
2933  
2934  
2935  
2936  
2937  
2938  
2939  
2940  
2941  
2942  
2943  
2944  
2945  
2946  
2947  
2948  
2949  
2950  
2951  
2952  
2953  
2954  
2955  
2956  
2957  
2958  
2959  
2960  
2961  
2962  
2963  
2964  
2965  
2966  
2967  
2968  
2969  
2970  
2971  
2972  
2973  
2974  
2975  
2976  
2977  
2978  
2979  
2980  
2981  
2982  
2983  
2984  
2985  
2986  
2987  
2988  
2989  
2990  
2991  
2992  
2993  
2994  
2995  
2996  
2997  
2998  
2999  
3000  
3001  
3002  
3003  
3004  
3005  
3006  
3007  
3008  
3009  
3010  
3011  
3012  
3013  
3014  
3015  
3016  
3017  
3018  
3019  
3020  
3021  
3022  
3023  
3024  
3025  
3026  
3027  
3028  
3029  
3030  
3031  
3032  
3033  
3034  
3035  
3036  
3037  
3038  
3039  
304

QY 3146 TGCTGCTTACAGACTTTTCTCTAGAAGCCGCTCTGCGTTTACTCTGTGTTTCAAGGGA 3205  
Db 3097 CGTTCCTTGGCTGGACTTTTCTCTAGATGCTGCTGCCATTACTC-----CAAAGTGA 3149  
QY 3206 CTTTGTGTAATCAAAATCATCTCTCAAGCAGGAGGAGCTGATATGAATTTATG 3265  
Db 3150 CTTCTATAAAATCAAAACCTCTCTCCACAGCGGGAGAGCCAAATAGAGACTTTGTTGG 3209  
QY 3266 GAGCATTGATCTGCATCAAGGCGCTTCTCAGCGCGCTTGAGTGAATGTGTACCTGAAG 3325  
Db 3210 TGAGCCGCTACCTCTGGGGCTTTTCCAGAGCTTGAGGGAAGCCATGTATCTGAA 3269  
QY 3326 TACAGTATATCTTGTGTAATACATAAAACAAA-----AGCATTTTGTGAAGGAAGC 3378  
Db 3270 TATAGTATATCTTGTGTAATACGTGAACCAACCAACCCGTTTTTGTCTAAGGGAAGC 3329  
QY 3379 TAATATGATTTTTT--AGTCTATGTTTAAATAATATGTAATTTTTCAGCTATTAG 3436  
Db 3330 TAAATATGATTTTTTAAATAATCTATGTTTTTAAATAATCTATGTAATTTTTCATCTATTAG 3389  
QY 3437 TGATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAAAATAAAATAAAATAAA 3496  
Db 3390 TGATATATTTTATGATGGAATAAACTTTCTACTGTAATAAAATAAAATAAAATAAA 3449  
QY 3497 AAAA 3500  
Db 3450 AAAA 3453

RESULT 6  
US-09-796-692-7538  
; Sequence 7538, Application US/09796692  
; Publication No. US20020198362A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES  
; FILE REFERENCE: 2077.001200  
; CURRENT APPLICATION NUMBER: US/09/796,692  
; CURRENT FILING DATE: 2001-03-01  
; PRIOR APPLICATION NUMBER: 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: 60/223,416  
; PRIOR FILING DATE: 2000-08-04  
; PRIOR APPLICATION NUMBER: 60/223,378  
; PRIOR FILING DATE: 2000-08-07  
; NUMBER OF SEQ ID NOS: 9597  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO: 7538  
; LENGTH: 645  
; TYPE: DNA  
; ORGANISM: Homo sapiens

US-09-796-692-7538

Query Match 18.1%; Score 633; DB 10; Length 645;  
Best Local Similarity 99.8%; Pred. No. 3.6e-150;  
Matches 644; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 2059 CTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGCGTGACACTGTTCAGAGCAAT 2118  
Db 1 CTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGCGTGACACTGTTCAGAGCAAT 60  
QY 2119 TACTTGAATTTTGAATCTGTTGCTATGTTGATCTTCTCAACTATCTTAAGAAGTAAAAA 2178  
Db 61 TACTTGAATTTTGAATCTGTTGCTATGTTGATCTTCTCAACTATCTTAAGAAGTAAAAA 120  
QY 2179 GAAAAATTTTCAGAGCTTGGACAGAGATTTTCAAGGAACACAAATTCATGTTTTTACCC 2238  
Db 121 GAAAAATTTTCAGAGCTTGGACAGAGATTTTCAAGGAACACAAATTTTCAGTTTTTACCC 180  
QY 2239 ACTTTCCATCATCTCCAAATTCAGAGCTGCTGCTTCAAGAGAAATTCAGATACACCC 2298  
Db 181 ACTTTCCATCATCTCCAAATTTCCAGCATGCTGCTTCAAGAGAAATTCAGATACACCC 240  
QY 2299 GACTCGGATCAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAGATGAAATGAA 2358  
Db 241 GACTCGGATCAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAGATGAAATGAA 300  
QY 2359 TATGAAAAACCAAAAAAGGCTGGAAGAAGAGGAGGACTTGAATGTGCTTACATTTGAAG 2418  
Db 301 TATGAAAAACCAAAAAAGGCTGGAAGAAGAGGAGGACTTGAATGTGCTTACATTTGAAG 360  
QY 2419 CTTCTTTGCTTTGCAATCAAGTTGCCAAAGAAATGGAATTTCTGGAATTTAAGTCGTGT 2478  
Db 361 CTTCTTTGCTTTGCAATCAAGTTGCCAAAGAAATGGAATTTCTGGAATTTAAGTCGTGT 420  
QY 2479 GTTCACAGAGACTGCGCCGACAGGAACTGCTTGTCAACCCACGGGAAAGTGGTGAAGATA 2538  
Db 421 GTTCACAGAGACTGCGCCGACAGGAACTGCTTGTCAACCCACGGGAAAGTGGTGAAGATA 480  
QY 2539 TGTGACTTTGGATTTGGCTCGAGATATCATGAGTGATTTCCAACTATGTTCTCAGGGCAAT 2598  
Db 481 TGTGACTTTGGATTTGGCTCGAGATATCATGAGTGATTTCCAACTATGTTCTCAGGGCAAT 540  
QY 2599 GCCGCTGCTGCTGTAATGGAATGGCCCGGAAAGGCTG-TTTGAAGGATCTACACCAT 2657  
Db 541 GCCGCTGCTGCTGTAATGGAATGGCCCGGAAAGGCTGTTTGAAGGATCTACACCAT 600  
QY 2658 TAAGATGATCTCTGCTCATATGGAATATTACTGTGGGAATCTT 2702  
Db 601 TAAGATGATCTCTGCTCATATGGAATATTACTGTGGGAATCTT 645

RESULT 7  
US-10-040-862-7538  
; Sequence 7538, Application US/10040862  
; Publication No. US20030078396A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Thera  
; TITLE OF INVENTION: Hematological Malignancies  
; FILE REFERENCE: 014058-013520US  
; CURRENT APPLICATION NUMBER: US/10/040,862  
; CURRENT FILING DATE: 2001-11-06  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303

;; PRIOR FILING DATE: 2000-04-28  
;; PRIOR APPLICATION NUMBER: US 60/200,779  
;; PRIOR FILING DATE: 2000-04-28  
;; PRIOR APPLICATION NUMBER: US 60/200,999  
;; PRIOR FILING DATE: 2000-05-01  
;; PRIOR APPLICATION NUMBER: US 60/202,084  
;; PRIOR FILING DATE: 2000-05-04  
;; PRIOR APPLICATION NUMBER: US 60/206,201  
;; PRIOR FILING DATE: 2000-05-22  
;; PRIOR APPLICATION NUMBER: US 60/218,950  
;; PRIOR FILING DATE: 2000-07-14  
;; PRIOR APPLICATION NUMBER: US 60/222,903  
;; PRIOR FILING DATE: 2000-08-03  
;; PRIOR APPLICATION NUMBER: US 60/223,416  
;; PRIOR FILING DATE: 2000-08-04  
;; PRIOR APPLICATION NUMBER: US 60/223,378  
;; PRIOR FILING DATE: 2000-08-07  
;; PRIOR APPLICATION NUMBER: US 09/796,692  
;; PRIOR FILING DATE: 2001-03-01  
;; NUMBER OF SEQ ID NOS: 10467  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 7538  
;; LENGTH: 645  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-040-862-7538

Query Match  
Best Local Similarity 18.1%; Score 633; DB 14; Length 645;  
Matches 644; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
QY 2059 CTGGGAAGCCAGAGATATTGTAACCTGCTGGGGCGTGACACTGTCAGGACCAATT 2118  
DB 1 CTGGGAAGCCAGAGATATTGTAACCTGCTGGGGCGTGACACTGTCAGGACCAATT 60  
QY 2119 TACTTGATTTTGAATCTGCTATGCTGATCTTCTCAACTATCTAAGAAGTAAAGA 2178  
DB 61 TACTTGATTTTGAATCTGCTATGCTGATCTTCTCAACTATCTAAGAAGTAAAGA 120  
QY 2179 GAAAAATTCACAGACTTGGACAGAGATTTCAAGGAACACATTTTCAAGTAAAGTAA 2238  
DB 121 GAAAAATTCACAGACTTGGACAGAGATTTTCAAGGAACACATTTTCAAGTAAAGTAA 180  
QY 2239 ACTTTCCAAATCACAATTCAGGACCTGCTGAGCATGCTGTTCAAGGAAGTAAAGTAA 2298  
DB 181 ACTTTCCAAATCACAATTCAGGACCTGCTGAGCATGCTGTTCAAGGAAGTAAAGTAA 240  
QY 2299 GACTCGGATCAAAATCTCAGGCGCTTCATGGGAATTCATTTCACTCTGAAGATGAATGAA 2358  
DB 541 GACTCGGATCAAAATCTCAGGCGCTTCATGGGAATTCATTTCACTCTGAAGATGAATGAA 600  
QY 2658 TAAGAGTGATCTCTGGTTCATATGGAATATTACTGTGGGAATCTTT 2702

DB 601 TAAGAGTGATCTCTGGTCAATATGGAATATTACTGTGGGAATCTTT 645

## RESULT 8

US-09-796-692-8085  
;; Sequence 8085, Application US/09796692  
;; Publication No. US20020198362A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Gaiger, Alexander  
;; APPLICANT: Algate, Paul A.  
;; APPLICANT: Mannion, Jane  
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERA  
;; FILE REFERENCE: 2077.001200  
;; CURRENT APPLICATION NUMBER: US/09/796,692  
;; CURRENT FILING DATE: 2001-03-01  
;; PRIOR APPLICATION NUMBER: 60/186,126  
;; PRIOR FILING DATE: 2000-03-01  
;; PRIOR APPLICATION NUMBER: 60/190,479  
;; PRIOR FILING DATE: 2000-03-17  
;; PRIOR APPLICATION NUMBER: 60/200,545  
;; PRIOR FILING DATE: 2000-04-27  
;; PRIOR APPLICATION NUMBER: 60/200,303  
;; PRIOR FILING DATE: 2000-04-28  
;; PRIOR APPLICATION NUMBER: 60/200,779  
;; PRIOR FILING DATE: 2000-04-28  
;; PRIOR APPLICATION NUMBER: 60/206,201  
;; PRIOR FILING DATE: 2000-05-22  
;; PRIOR APPLICATION NUMBER: 60/218,950  
;; PRIOR FILING DATE: 2000-07-14  
;; PRIOR APPLICATION NUMBER: 60/222,903  
;; PRIOR FILING DATE: 2000-08-03  
;; PRIOR APPLICATION NUMBER: 60/223,416  
;; PRIOR FILING DATE: 2000-08-04  
;; PRIOR APPLICATION NUMBER: 60/223,378  
;; NUMBER OF SEQ ID NOS: 9597  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 8085  
;; LENGTH: 610  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-09-796-692-8085

Query Match  
Best Local Similarity 17.4%; Score 608.4; DB 10; Length 610;  
Matches 609; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2059 CTGGGAAGCCAGAGATATTGTAACCTGCTGGGGCGTGACACTGTCAGGACCAATT 2118  
DB 1 CTGGGAAGCCAGAGATATTGTAACCTGCTGGGGCGTGACACTGTCAGGACCAATT 60  
QY 2119 TACTTGATTTTGAATCTGCTATGCTGATCTTCTCAACTATCTAAGAAGTAAAGA 2178  
DB 61 TACTTGATTTTGAATCTGCTATGCTGATCTTCTCAACTATCTAAGAAGTAAAGA 120  
QY 2179 GAAAAATTCACAGACTTGGACAGAGATTTTCAAGGAACACATTTTCAAGTAAAGTAA 2238  
DB 121 GAAAAATTCACAGACTTGGACAGAGATTTTCAAGGAACACATTTTCAAGTAAAGTAA 180  
QY 2239 ACTTTCCAAATCACAATTCAGGACCTGCTGAGCATGCTGTTCAAGGAAGTAAAGTAA 2298  
DB 181 ACTTTCCAAATCACAATTCAGGACCTGCTGAGCATGCTGTTCAAGGAAGTAAAGTAA 240  
QY 2299 GACTCGGATCAAAATCTCAGGCGCTTCATGGGAATTCATTTCACTCTGAAGATGAATGAA 2358  
DB 241 GACTCGGATCAAAATCTCAGGCGCTTCATGGGAATTCATTTCACTCTGAAGATGAATGAA 300  
QY 2359 TATGAAAAACCAAAAGGCTGGAAGAGAGAGGAGGACTTGAATGTGCTTACATTTGAAGAT 2418

Db 301 TATGAAACCAAAAGGCTGGAAGAGAGGAGGACTTGAATGTGCTTACATTTGAAGAT 360  
Qy 2419 CTTCTTTGCTTTGTCATATCAAGTTGCCAAAGGAATGGAATTTCTGGAATTTAAAGTCGTGT 2478  
Db 361 CTTCTTTGCTTTGTCATATCAAGTTGCCAAAGGAATGGAATTTCTGGAATTTAAAGTCGTGT 420  
Qy 2479 GTTTCACAGAGACCTGGCCGAGGAACTGCTTTGTCAACCCGAGGAAAGTGTGAAGATA 2538  
Db 421 GTTTCACAGAGACCTGGCCGAGGAACTGCTTTGTCAACCCGAGGAAAGTGTGAAGATA 480  
Qy 2539 TGTGACTTTGATGGCTCGAGATATCATGATGATGATTCCAACTATGTTGTGAGGGCAAT 2598  
Db 481 TGTGACTTTGATGGCTCGATATATCATGATGATTCCAACTATGTTGTGAGGGCAAT 540  
Qy 2599 GCCCGTCTGCTGTAATAATGATGTCGCCCGAGGAACTGTTTGAAGGCAATCTACACCAAT 2658  
Db 541 GCCCGTCTGCTGTAATAATGATGTCGCCCGAGGAACTGTTTGAAGGCAATCTACACCAAT 600  
Qy 2659 AAGAGTGATG 2668  
Db 601 AAGAGTGATG 610

RESULT 9  
US-10-040-862-8085  
; Sequence 8085, Application US/10040862  
; Publication No. US20030078396A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc  
; TITLE OF INVENTION: Hematological Malignancies  
; TITLE OF INVENTION: Hematological Malignancies  
; FILE REFERENCE: 014058-01352005  
; CURRENT APPLICATION NUMBER: US/10/040,862  
; CURRENT FILING DATE: 2001-11-06  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: US 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: US 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: US 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: US 60/223,416  
; PRIOR FILING DATE: 2000-08-04  
; PRIOR APPLICATION NUMBER: US 60/223,378  
; PRIOR FILING DATE: 2000-08-07  
; PRIOR APPLICATION NUMBER: US 09/796,692  
; PRIOR FILING DATE: 2001-03-01  
; NUMBER OF SEQ ID NOS: 10467  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8085  
; LENGTH: 610  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-040-862-8085

Query Match 17.4%; Score 608.4; DB 14; Length 610;  
Best Local Similarity 99.8%; Pred. No. 6.2e-144; Indels 0; Gaps 0;  
Matches 609; Conservative 0; Mismatches 1;  
Qy 2059 CTGGGAAGCCACGAGAAATATTTGTGAACCTGCTGGGGCGTGACACTGTTCAGGACCAAT 2118  
Db 1 CTGGGAAGCCACGAGAAATATTTGTGAACCTGCTGGGGCGTGACACTGTTCAGGACCAAT 60  
Qy 2119 TACTTGATTTTGAATACTGTGTGATGATCTTCTCAACTATCTTAAGAAATGAAAGA 2178  
Db 61 TACTTGATTTTGAATACTGTGTGATGATCTTCTCAACTATCTTAAGAAATGAAAGA 120  
Qy 2179 GAAAAATTTTCAGGACTTGCACAGAGATTTTCAAGGAACAAATTTTCAGTTTTTACCCC 2238  
Db 121 GAAAAATTTTCAGGACTTGCACAGAGATTTTCAAGGAACAAATTTTCAGTTTTTACCCC 180  
Qy 2239 ACTTTCCAATCACATCCAAATTTCCAGCATGCTGGTTCAAGAGAAATTCAGATACACCCG 2298  
Db 181 ACTTTCCAATCACATCCAAATTTCCAGCATGCTGGTTCAAGAGAAATTCAGATACACCCG 240  
Qy 2299 GACTCGGATCAAAATCTCAGGGCTTCAAGGAAATTCATTTCACTCTGAAATGAAATGAA 2358  
Db 241 GACTCGGATCAAAATCTCAGGGCTTCAAGGAAATTCATTTCACTCTGAAATGAAATGAA 300  
Qy 2359 TATGAAACCAAAAGGCTGGAAGAGAGGAGGACTTGAATGTGCTTACATTTGAAGAT 2418  
Db 301 TATGAAACCAAAAGGCTGGAAGAGAGGAGGACTTGAATGTGCTTACATTTGAAGAT 360  
Qy 2419 CTTCTTTGCTTTGTCATATCAAGTTTCCAAAGAAATGGAATTTCTGAAATTTAAAGTCGTGT 2478  
Db 361 CTTCTTTGCTTTGTCATATCAAGTTTCCAAAGAAATGGAATTTCTGAAATTTAAAGTCGTGT 420  
Qy 2479 GTTTCACAGAGACCTGGCCGAGGAACTGCTTGTCAACCCGAGGAAAGTGTGAAAGATA 2538  
Db 421 GTTTCACAGAGACCTGGCCGAGGAACTGCTTGTCAACCCGAGGAAAGTGTGAAAGATA 480  
Qy 2539 TGTGACTTTGATGGCTCGAGATATCATGATGATTCCAACTATGTTGTTCAGGGCAAT 2598  
Db 481 TGTGACTTTGATGGCTCGATATATCATGATGATTCCAACTATGTTGTTCAGGGCAAT 540  
Qy 2599 GCCCGTCTGCTGTAATAATGATGTCGCCCGAGGAACTGTTTGAAGGCAATCTACACCAAT 2658  
Db 541 GCCCGTCTGCTGTAATAATGATGTCGCCCGAGGAACTGTTTGAAGGCAATCTACACCAAT 600  
Qy 2659 AAGAGTGATG 2668  
Db 601 AAGAGTGATG 610  
RESULT 10  
US-09-796-692-8132/c  
; Sequence 8132, Application US/09796692  
; Publication No. US20020198362A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERA  
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES  
; FILE REFERENCE: 2077.001200  
; CURRENT APPLICATION NUMBER: US/09/796,692  
; CURRENT FILING DATE: 2001-03-01  
; PRIOR APPLICATION NUMBER: 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,999





QY 2648 TCTACACCAATTAAGAGTGATGCTGGTGCATATGGAATATTACTGTGGGAAATCTTCTCAC 2707  
Db 248 TCTACACCAATTAAGAGTGATGCTGGTGCATATGGAATATTACTGTGGGAAATCTTCTCAC 189  
QY 2708 TTGGTGTGAATCCCTTACCTGGCATTCGGTTGATGCTTAACCTTCTACAACTGATTCAAA 2767  
Db 188 TTGGTGTGAATCTTTACCTGGCATTCGGTTGATGCTTAACCTTCTACAACTGATTCAAA 129  
QY 2768 ATGGAATTTAAATGATCAGACCAATTTTATGCTACAGAGAAATATACATTATTAATGCAAT 2827  
Db 128 ATGGAATTTAAATGATCAGACCAATTTTATGCTACAGAGAAATATACATTATTAATGCAAT 69  
QY 2828 CCTGCTGGGCTTTTGACTCAAGGAAACGGCCATCCTTCCCTAAATTTGACTTCGTTTTAG 2887  
Db 68 CCTGCTGGGCTTTTGACTCAAGGAAACGGCCATCCTTCCCTAAATTTGACTTCGTTTTAG 9  
QY 2888 GATGTCAG 2895  
Db 8 GATGTCAG 1

RESULT 12  
US-09-796-692-8927  
; Sequence 8927, Application US/09796692  
; Publication No. US20020198362A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
; FILE REFERENCE: 2077.001200  
; CURRENT APPLICATION NUMBER: US/09/796,692  
; CURRENT FILING DATE: 2001-03-01  
; PRIOR APPLICATION NUMBER: 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,779  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 60/202,084  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: 60/223,416  
; PRIOR FILING DATE: 2000-08-04  
; PRIOR APPLICATION NUMBER: 60/223,378  
; PRIOR FILING DATE: 2000-08-07  
; NUMBER OF SEQ ID NOS: 9597  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8927  
; LENGTH: 605  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-796-692-8927

Query Match 17.3%; Score 605; DB 10; Length 605;  
Best Local Similarity 100.0%; Pred. No. 4.5e-143;  
Matches 605; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2059 CTGGAAGCCAGAGAATTTGAACTCTGGGGGGTGCACACTGTGACGACCAAT 2118  
,|

Db 1 CTGGAAGCCAGAGAATTTGTAACCTGCTGGGGGGTGCACACTGTGACGACCAAT 60  
QY 2119 TACTTGATTTTGAATACTGTTGCTATGGTGATCTTCTCAAATCTAAGAAAGTAAAGA 2178  
Db 61 TACTTGATTTTGAATACTGTTGCTATGGTGATCTTCTCAAATCTAAGAAAGTAAAGA 120  
QY 2179 GAAAAATTTACAGAGACTTGGACAGAGATTTTCAAGGAAACAAATTTTACAGTTTACCCC 2238  
Db 121 GAAAAATTTACAGAGACTTGGACAGAGATTTTCAAGGAAACAAATTTTACAGTTTACCCC 180  
QY 2239 ACTTTTCAATCACAATCCAAATTCAGGCTTCAAGGAAAGTTCAGATACACCCG 2298  
Db 181 ACTTTTCAATCACAATCCAAATTCAGGCTTCAAGGAAAGTTCAGATACACCCG 240  
QY 2299 GACTCGGATCAAAATCTCAGGGCTTTCAGGGAAATTCATTCTCTGAAGATGAAATTTGAA 2358  
Db 241 GACTCGGATCAAAATCTCAGGGCTTTCAGGGAAATTCATTCTCTGAAGATGAAATTTGAA 300  
QY 2359 TATGAAAACCAAAAAAGGCTGGAAGAGAGAGAGACTTGAATGTGCTTACATTGGAAGAT 2418  
Db 301 TATGAAAACCAAAAAAGGCTGGAAGAGAGAGAGAGACTTGAATGTGCTTACATTGGAAGAT 360  
QY 2419 CTCTCTTGTGCTATCAAGTTTGCACAAAGAAATGGAATTTCTGGAATTTTAAAGTCGTGT 2478  
Db 361 CTCTCTTGTGCTATCAAGTTTGCACAAAGAAATGGAATTTCTGGAATTTTAAAGTCGTGT 420  
QY 2479 GTTCAAGAGACCTGGCGCCAGGAAACGTGCTTGTCAACCCAGGAAAGTGTGGAAGATA 2538  
Db 421 GTTCAAGAGACCTGGCGCCAGGAAACGTGCTTGTCAACCCAGGAAAGTGTGGAAGATA 480  
QY 2539 TGTGACTTTGATTTGGCTGCGAGATATCATAGTGAATTCACATATGTTGTCTAGGGGCAAT 2598  
Db 481 TGTGACTTTGATTTGGCTGCGAGATATCATAGTGAATTCACATATGTTGTCTAGGGGCAAT 540  
QY 2599 GCCGCTCTGCTGTAAATGATGCTGCCCCGAAAGCTGTTTGAAGGATCTTACACCAAT 2658  
Db 541 GCCGCTCTGCTGTAAATGATGCTGCCCCGAAAGCTGTTTGAAGGATCTTACACCAAT 600  
QY 2659 AAGAG 2663  
Db 601 AAGAG 605

RESULT 13  
US-10-040-862-8927  
; Sequence 8927, Application US/10040862  
; Publication No. US20030078396A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therap  
; FILE REFERENCE: 014058-013520US  
; CURRENT APPLICATION NUMBER: US/10/040,862  
; CURRENT FILING DATE: 2001-11-05  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: US 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: US 60/206,201

;; PRIOR FILING DATE: 2000-05-22  
;; PRIOR APPLICATION NUMBER: US 60/218,950  
;; PRIOR FILING DATE: 2000-07-14  
;; PRIOR APPLICATION NUMBER: US 60/222,903  
;; PRIOR FILING DATE: 2000-08-03  
;; PRIOR APPLICATION NUMBER: US 60/223,416  
;; PRIOR FILING DATE: 2000-08-04  
;; PRIOR APPLICATION NUMBER: US 60/223,378  
;; PRIOR FILING DATE: 2000-08-07  
;; PRIOR APPLICATION NUMBER: US 09/796,692  
;; PRIOR FILING DATE: 2001-03-01  
;; NUMBER OF SEQ ID NOS: 10467  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 8927  
;; LENGTH: 605  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-040-862-8927

Query Match 17.3%; Score 605; DB 14; Length 605;  
Best Local Similarity 100.0%; Pred. No. 4.5e-143;  
Matches 605; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2059 CTGGAAGCCACGAGATATTGTGAACCTGCTGGGGCGTGACACCTGTGACGACCAATT 2118  
Db 1 CTGGAAGCCACGAGATATTGTGAACCTGCTGGGGCGTGACACCTGTGACGACCAATT 60  
Qy 2119 TACTTGATTTTGAATCTGTTGATGCTGATCTTCAACTATCTAAGAGTAAAGA 2178  
Db 61 TACTTGATTTTGAATCTGTTGATGCTGATCTTCAACTATCTAAGAGTAAAGA 120  
Qy 2179 GAAAAATTTTCAGGACTTGGACAGAGATTTTCAAGGAACACAAATTTACAGTTTACCCC 2238  
Db 121 GAAAAATTTTCAGGACTTGGACAGAGATTTTCAAGGAACACAAATTTACAGTTTACCCC 180  
Qy 2239 ACTTTCCAAATCATCCAAATTCAGCATGCTGTTCAAGAGAAATTCAGATACACCCG 2298  
Db 181 ACTTTCCAAATCATCCAAATTCAGCATGCTGTTCAAGAGAAATTCAGATACACCCG 240  
Qy 2299 GACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAGATGAATGAA 2358  
Db 241 GACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAGATGAATGAA 300  
Qy 2359 TATGAAAAACCAAAAAAGCGCTGGAAGAGAGAGGACTTGAATGCTTACATTTCAAGAT 2418  
Db 301 TATGAAAAACCAAAAAAGCGCTGGAAGAGAGAGGACTTGAATGCTTACATTTCAAGAT 360  
Qy 2419 CTTCTTTGCTTTGCATATCAAGTTTGCACCAAGGAATGGAATTTCTGGAATTTAAGTCGTGT 2478  
Db 361 CTTCTTTGCTTTGCATATCAAGTTTGCACCAAGGAATGGAATTTCTGGAATTTAAGTCGTGT 420  
Qy 2479 GTTTCACAGAGACCTGGCCGCCAGGNA CGTGCTTGTTCACCCACGGGAAAGTGGTGAAGATA 2538  
Db 421 GTTTCACAGAGACCTGGCCGCCAGGNA CGTGCTTGTTCACCCACGGGAAAGTGGTGAAGATA 480  
Qy 2539 TGTGACTTTTGAATGGCTCGAGATATCATGATGATTCCTCAACTATGTTGTGAGGGCAAT 2598  
Db 481 TGTGACTTTTGAATGGCTCGAGATATCATGATGATTCCTCAACTATGTTGTGAGGGCAAT 540  
Qy 2599 GCCCGTCTGCCGTGTAATAATGATGGCCGCCGAAAGCCCTGTTTGAAGGCAATCTACACCAAT 2658  
Db 541 GCCCGTCTGCCGTGTAATAATGATGGCCGCCGAAAGCCCTGTTTGAAGGCAATCTACACCAAT 600  
Qy 2659 AAGAG 2663  
Db 601 AAGAG 605

## RESULT 14

US-09-796-692-8273  
; Sequence 8273, Application US/09796692  
; Publication No. US20020198362A1  
; GENERAL INFORMATION:

;; APPLICANT: Gaiger, Alexander  
;; APPLICANT: Algate, Paul A.  
;; APPLICANT: Mannion, Jane  
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERA  
;; FILE REFERENCE: 2077.001200  
;; CURRENT APPLICATION NUMBER: US/09/796,692  
;; PRIOR FILING DATE: 2001-03-01  
;; PRIOR APPLICATION NUMBER: 60/186,126  
;; PRIOR FILING DATE: 2000-03-01  
;; PRIOR APPLICATION NUMBER: 60/190,479  
;; PRIOR FILING DATE: 2000-03-17  
;; PRIOR APPLICATION NUMBER: 60/200,545  
;; PRIOR FILING DATE: 2000-04-27  
;; PRIOR APPLICATION NUMBER: 60/200,303  
;; PRIOR FILING DATE: 2000-04-28  
;; PRIOR APPLICATION NUMBER: 60/200,779  
;; PRIOR FILING DATE: 2000-04-28  
;; PRIOR APPLICATION NUMBER: 60/200,999  
;; PRIOR FILING DATE: 2000-05-01  
;; PRIOR APPLICATION NUMBER: 60/202,084  
;; PRIOR FILING DATE: 2000-05-04  
;; PRIOR APPLICATION NUMBER: 60/206,201  
;; PRIOR FILING DATE: 2000-05-22  
;; PRIOR APPLICATION NUMBER: 60/218,950  
;; PRIOR FILING DATE: 2000-07-14  
;; PRIOR APPLICATION NUMBER: 60/222,903  
;; PRIOR FILING DATE: 2000-08-03  
;; PRIOR APPLICATION NUMBER: 60/223,416  
;; PRIOR FILING DATE: 2000-08-04  
;; PRIOR APPLICATION NUMBER: 60/223,378  
;; PRIOR FILING DATE: 2000-08-07  
;; NUMBER OF SEQ ID NOS: 9597  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 8273  
;; LENGTH: 601  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; NAME/KEY: unsure  
;; LOCATION: (586)  
;; OTHER INFORMATION: n=A, T, C or G  
US-09-796-692-8273

Query Match 17.1%; Score 598.4; DB 10; Length 601;  
Best Local Similarity 99.7%; Pred. No. 2.1e-141;  
Matches 599; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 2059 CTGGAAGCCACGAGATATTGTGAACCTGCTGGGGCGTGACACCTGTGACGACCAATT 2118  
Db 1 CTGGAAGCCACGAGATATTGTGAACCTGCTGGGGCGTGACACCTGTGACGACCAATT 60  
Qy 2119 TACTTGATTTTGAATCTGTTGATGCTTCTCAACTATCTAAGAGTAAAGA 2178  
Db 61 TACTTGATTTTGAATCTGTTGATGCTTCTCAACTATCTAAGAGTAAAGA 120  
Qy 2179 GAAAAATTTTCAGGACTTGGACAGAGATTTTCAAGGAACACAAATTTACAGTTTACCCC 2238  
Db 121 GAAAAATTTTCAGGACTTGGACAGAGATTTTCAAGGAACACAAATTTACAGTTTACCCC 180  
Qy 2239 ACTTTCAATCATCATCAATTCAGCATGCTGTTTCAAGAGAAATTCAGATACACCCG 2298  
Db 181 ACTTTCAATCATCATCAATTCAGCATGCTGTTTCAAGAGAAATTCAGATACACCCG 240  
Qy 2299 GACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAGATGAATGAA 2358  
Db 241 GACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAGATGAATGAA 300  
Qy 2359 TATGAAAAACCAAAAAAGCGCTGGAAGAGAGAGGACTTGAATGCTTACATTTCAAGAT 2418  
Db 301 TATGAAAAACCAAAAAAGCGCTGGAAGAGAGAGGACTTGAATGCTTACATTTCAAGAT 360  
Qy 2419 CTTCTTTGCTTTGCATATCAAGTTTGCACCAAGGAATGGAATTTCTGGAATTTAAGTCGTGT 2478

||||| 361 CTCTCTTGTCTTGCATCAAGTTGCCAAAGGAATGGAATTTCTGGAATTTAAAGTCGTGT 420  
QY 2479 GTTCACAGAGACCTGGCGCCAGGAACGTGCTTGTCTACCCACGGGAAAGTGGTGAAGATA 2538  
Db 421 GTTCACAGAGACCTGGCGCCAGGAACGTGCTTGTCTACCCACGGGAAAGTGGTGAAGATA 480  
QY 2539 TGTGACTTTGGATTGGCTCGAGATATCATGAGTGATTTCCAACTATGTTGTCAGGGGCAAT 2598  
Db 481 TGTGACTTTGGATTGGCTCGATATATCATGAGTGATTTCCAACTATGTTGTCAGGGGCAAT 540  
QY 2599 GCCCGTCTGCTGTAAATGAGTGGCCCGGAAAGCCTGTTGGAAGCATCTACACCAT 2658  
Db 541 GCCCGTCTGCTGTAAATGAGTGGCCCGGAAAGCCTGTTGGAAGCATCTACACCAT 600  
QY 2659 A 2659  
Db 601 A 601

RESULT 15  
US-10-040-862-8273  
; Sequence 8273, Application US/10040862  
; Publication No. US20030078396A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; FILE REFERENCE: 014058-013520US  
; CURRENT APPLICATION NUMBER: US/10/040,862  
; CURRENT FILING DATE: 2001-11-06  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: US 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: US 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: US 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: US 60/223,416  
; PRIOR FILING DATE: 2000-08-04  
; PRIOR APPLICATION NUMBER: US 60/223,378  
; PRIOR FILING DATE: 2000-08-07  
; PRIOR APPLICATION NUMBER: US 09/796,692  
; PRIOR FILING DATE: 2001-03-01  
; NUMBER OF SEQ ID NOS: 10467  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8273  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (586)  
; OTHER INFORMATION: n=A,T,C or G  
US-10-040-862-8273

Query Match 17.1%; Score 598.4; DB 14; Length 601;  
Best Local Similarity 99.7%; Pred. No. 2.1e-141;  
Matches 599; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 2059 CTGGGAAGCCACAGAGAAATATTGTGAACCTCTCGGGGCGTGCACACTGTGAGGACCAATT 2118  
Db 1 CTGGGAAGCCACAGAGAAATATTGTGAACCTCTCGGGGCGTGCACACTGTGAGGACCAATT 60  
QY 2119 TACTTGATTTTGAATACCTGTTGCTATGTTGATCTCTCTCAACTATCTTAAGAGTAAAGA 2178  
Db 61 TACTTGATTTTGAATACCTGTTGCTATGTTGATCTCTCTCAACTATCTTAAGAGTAAAGA 120  
QY 2179 GAAAAATTTTCCACAGGACTTGGACAGAGATTTTCAAGGAAACACAATTTTTCAGTTTTTACCCC 2238  
Db 121 GAAAAATTTTCCACAGGACTTGGACAGAGATTTTCAAGGAAACACAATTTTTCAGTTTTTACCCC 180  
QY 2239 ACTTTTCCAATCACATCCAAATTTCCAGCATGCTGGTTCAAGAGAAAGTTTCAGATACACCG 2298  
Db 181 ACTTTTCCAATCACATCCAAATTTCCAGCATGCTGGTTCAAGAGAAAGTTTCAGATACACCG 240  
QY 2299 GACTCGGATCAAAATCTCAGGCTTTCAGGGAAATTCATTTCACTCTGAAGATGAAATTGAA 2358  
Db 241 GACTCGGATCAAAATCTCAGGCTTTCAGGGAAATTCATTTCACTCTGAAGATGAAATTGAA 300  
QY 2359 TATGAAAAACCAAAAAGGCTGGAAGAGAGAGGACTTGAATGTGCTTACATTTTGAAGAT 2418  
Db 301 TATGAAAAACCAAAAAGGCTGGAAGAGAGAGGACTTGAATGTGCTTACATTTTGAAGAT 360  
QY 2419 CTCTTTTGTCTTGCATATCAAGTTGCCAAAGGAATGGAAATTTCTTGGAAATTTAAGTCGTGT 2478  
Db 361 CTCTTTTGTCTTGCATATCAAGTTGCCAAAGGAATGGAAATTTCTTGGAAATTTAAGTCGTGT 420  
QY 2479 GTTCACAGAGACCTGGCGCCAGGAAACGTGCTTGTCTACCCACGGGAAAGTGGTGAAGATA 2538  
Db 421 GTTCACAGAGACCTGGCGCCAGGAAACGTGCTTGTCTACCCACGGGAAAGTGGTGAAGATA 480  
QY 2539 TGTGACTTTGGATTGGCTCGAGATATCATGAGTGATTTCCAACTATGTTTGCAGGGGCAAT 2598  
Db 481 TGTGACTTTGGATTGGCTCGAGATATCATGAGTGATTTCCAACTATGTTTGCAGGGGCAAT 540  
QY 2599 GCCCGTCTGCTGTAAATGAGTGGCCCGGAAAGCCTGTTTGAAGGATCTTACACCAT 2658  
Db 541 GCCCGTCTGCTGTAAATGAGTGGCCCGGAAAGCCTGTTTGAAGGATCTTACACCAT 600  
QY 2659 A 2659  
Db 601 A 601

Search completed: August 28, 2003, 13:26:40  
Job time : 1027.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 15:21:07 ; Search time 284.45 Seconds  
(without alignments)  
5432.528 Million cell updates/sec

Title: US-09-919-408A-3

Perfect score: 3501

Sequence: 1 CGAGCGCGCATCCGAGGCT.....AAAAAAAAAAAAAAAAAAAA 3501

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:  
5: /cgn2\_6/ptodata/2/ina/PTUS\_COMB.seq:  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3501	100.0	3501	1	US-07-977-451-3
2	3501	100.0	3501	1	US-08-252-517-3
3	3501	100.0	3501	1	US-07-906-397A-3
4	3501	100.0	3501	1	US-08-601-891-3
5	3501	100.0	3501	2	US-09-021-324-3
6	3501	100.0	3501	5	PCT-US92-05401-3
7	3501	100.0	3501	5	PCT-US92-09893-3
8	3475	99.3	3475	1	US-08-222-299-3
9	3475	99.3	3475	2	US-08-434-878-3
10	3475	99.3	3475	5	PCT-US95-03718-3
11	3474.4	99.2	3476	1	US-08-183-211-1
12	3474.4	99.2	3476	5	PCT-US95-00176A-1
13	3083.2	88.1	3120	4	US-08-222-616-22
14	3083.2	88.1	3120	4	US-08-446-648-22
15	3083.2	88.1	3120	5	PCT-US95-04228-22
16	2274.6	65.0	3521	1	US-08-222-299-1
17	2274.6	65.0	3521	2	US-08-434-878-1
18	2274.6	65.0	3521	5	PCT-US95-03718-1
19	2274.6	65.0	3521	5	PCT-US95-03718-1
20	2227.2	63.6	3453	1	US-07-977-451-1
21	2227.2	63.6	3453	1	US-07-946-507-1
22	2227.2	63.6	3453	1	US-08-252-517-1
23	2227.2	63.6	3453	1	US-07-906-397A-1
24	2227.2	63.6	3453	1	US-08-601-891-1
25	2227.2	63.6	3453	2	US-09-021-324-1
26	2227.2	63.6	3453	5	PCT-US92-02750-1
27	2227.2	63.6	3453	5	PCT-US92-05401-1

28 2227.2 63.6 3453 5 PCT-US92-09893-1 Sequence 1, Appli  
29 624.4 17.8 1894 3 US-07-912-122-3 Sequence 3, Appli  
30 624.4 17.8 1894 5 PCT-US93-06404-3 Sequence 3, Appli  
31 332 9.5 332 5 PCT-US92-02750-3 Sequence 3, Appli  
32 313.2 8.9 5084 1 US-08-306-691B-21 Sequence 21, Appli  
33 313.2 8.9 5084 5 PCT-US93-06251-25 Sequence 25, Appli  
34 302.6 8.6 4054 1 US-08-180-195-35 Sequence 35, Appli  
35 302.6 8.6 4054 1 US-08-477-329-35 Sequence 35, Appli  
36 302.6 8.6 4054 2 US-08-475-458-35 Sequence 35, Appli  
37 302.6 8.6 4054 3 US-08-980-400-35 Sequence 35, Appli  
38 302.6 8.6 4054 3 US-09-583-459A-35 Sequence 35, Appli  
39 302.6 8.6 4054 3 US-09-583-210-35 Sequence 35, Appli  
40 302.6 8.6 4054 4 US-09-583-449A-35 Sequence 35, Appli  
41 302.6 8.6 4054 4 US-09-435-059-35 Sequence 35, Appli  
42 302.6 8.6 4100 1 US-08-168-917-3 Sequence 3, Appli  
43 302.6 8.6 4100 2 US-08-460-510-3 Sequence 3, Appli  
44 302.6 8.6 4100 2 US-08-460-490-3 Sequence 3, Appli  
45 302.6 8.6 4100 5 PCT-US92-00730-3 Sequence 3, Appli

## ALIGNMENTS

## RESULT 1

US-07-977-451-3  
; Sequence 3, Application US/07977451  
; Patent No. 5270458  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ImClone Systems Incorporated  
; STREET: 180 Varick Street  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/977,451  
; FILING DATE: 19921119  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US UNASSIGNED  
; FILING DATE: 12-NOV-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/906,397  
; FILING DATE: 26-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US PCT/US92/05401  
; FILING DATE: 26-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: TW 81102961  
; FILING DATE: 15-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US PCT/US92/02750  
; FILING DATE: 02-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/813,593  
; FILING DATE: 24-DEC-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/793,065  
; FILING DATE: 15-NOV-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/728,913  
; FILING DATE: 28-JUN-1991  
; PRIOR APPLICATION DATA:





```

CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/252,517
  FILING DATE: 31-OCT-1994
  CLASSIFICATION: 530
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 07/977,451
    FILING DATE: 19-NOV-1992
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 07/906,397
    FILING DATE: 26-JUN-1992
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US PCT/US92/05401
    FILING DATE: 26-JUN-1992
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: TW 81102961
    FILING DATE: 15-APR-1992
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US PCT/US92/02750
    FILING DATE: 02-APR-1992
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 07/813,593
    FILING DATE: 24-DEC-1991
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 07/793,065
    FILING DATE: 15-NOV-1991
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 07/728,913
    FILING DATE: 28-JUN-1991
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 07/679,666
    FILING DATE: 02-APR-1991
  ATTORNEY/AGENT INFORMATION:
    NAME: Feit, Irving N.
    REGISTRATION NUMBER: 28,601
    REFERENCE/DOCKET NUMBER: LEM-3-7P
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 212-645-1405
    TELEFAX: 212-645-2054
    INFORMATION FOR SEQ ID NO: 3:
      SEQUENCE CHARACTERISTICS:
        LENGTH: 3501 base pairs
        TYPE: nucleic acid
        STRANDEDNESS: double
        TOPOLOGY: linear
        MOLECULE TYPE: cDNA
        HYPOTHETICAL: NO
        ANTI-SENSE: NO
      FRAGMENT TYPE: N-terminal
      FEATURE:
        NAME/KEY: CDS
        LOCATION: 58..3039
      FEATURE:
        NAME/KEY: mat_peptide
        LOCATION: 139..3036
      FEATURE:
        NAME/KEY: sig_peptide
        LOCATION: 58..138
US-08-252-517-3

Query Match      100.0%; Score 3501; DB 1; Length 3501;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy 1 CGAGCGCGCATCCGAGGCGCTGGCGCGCCCTGGGGGACCCCGGGCTCCGGAGGCCCATG 60
Db 1 CGAGCGCGCATCCGAGGCGCTGGCGCGCCCTGGGGGACCCCGGGCTCCGGAGGCCCATG 60

Qy 61 CCGGCGTTGGCGCGCGACGCGCGGCACCGTGCCGCTGCTCGTTGTTTTTCTGCAATGATA 120
Db 61 CCGGCGTTGGCGCGCGACGCGCGGCACCGTGCCGCTGCTCGTTGTTTTTCTGCAATGATA 120

Qy 121 TTTGGGACATTATCAAAATCAAGATCGCTGTGATCACTGTTGTTTAAATCAATCATAG 180

```



QY 1261 AGCATATCAAAGTTTGGCAATCATTAAGCACACAGCCAGGAGAAATATATATTCATCGAGAA 1320  
DB 1261 AGCATATCAAAGTTTGGCAATCATTAAGCACACAGCCAGGAGAAATATATATTCATCGAGAA 1320  
QY 1321 AATGATGATGCCAAATTTACCAAAATGTTTACGCTGAATATTAAGAAAGAAACCTCAAGTG 1380  
DB 1321 AATGATGATGCCAAATTTACCAAAATGTTTACGCTGAATATTAAGAAAGAAACCTCAAGTG 1380  
QY 1381 CTCGAGAAAGCATCGGCAAGTCAGCGCTCCTGTTTCTCGATGATACCCATTCATCT 1440  
DB 1381 CTCGAGAAAGCATCGGCAAGTCAGCGCTCCTGTTTCTCGATGATACCCATTCATCT 1440  
QY 1441 TGGACCTGGAAAGATGTTTACCAAGTCTCCCACTGACACAGAGAGATCAAGAGGA 1500  
DB 1441 TGGACCTGGAAAGATGTTTACCAAGTCTCCCACTGACACAGAGAGATCAAGAGGA 1500  
QY 1501 GTCTGGAATAGAAAGCTTAACAGAAAGTGTGACAGTGGGTGTCGAGCAGTACTCTA 1560  
DB 1501 GTCTGGAATAGAAAGCTTAACAGAAAGTGTGACAGTGGGTGTCGAGCAGTACTCTA 1560  
QY 1561 AACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTGTCATACAAATTCCTTGGC 1620  
DB 1561 AACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTGTCATACAAATTCCTTGGC 1620  
QY 1621 ACATCTTGTGAGACGATCCTTTTAACTCTCCAGGCCCTTCCCTTTTCATCCAAAGCAAC 1680  
DB 1621 ACATCTTGTGAGACGATCCTTTTAACTCTCCAGGCCCTTCCCTTTTCATCCAAAGCAAC 1680  
QY 1681 ATCTCATTTCTATGCAACAAATTTGGTGTGTTCTCTCTCAATTTGTCGTTTAAACCTGCTA 1740  
DB 1681 ATCTCATTTCTATGCAACAAATTTGGTGTGTTCTCTCTCAATTTGTCGTTTAAACCTGCTA 1740  
QY 1741 ATTTCTCAAGTACAAAGCAATTTAGTGTAAAGCCAGCTACAGATGGTACAGGTG 1800  
DB 1741 ATTTCTCAAGTACAAAGCAATTTAGTGTAAAGCCAGCTACAGATGGTACAGGTG 1800  
QY 1801 ACCGGCTCTCAGATAATGAGTACTTCTACGTGATTTTCAGAGAAATGATGAATGATCTC 1860  
DB 1801 ACCGGCTCTCAGATAATGAGTACTTCTACGTGATTTTCAGAGAAATGATGAATGATCTC 1860  
QY 1861 AAATGGAGTTTCCAAAGAGAAATTTAGAGTTTGGGAAAGGTACTAGGATCAGGTGCTTTT 1920  
DB 1861 AAATGGAGTTTCCAAAGAGAAATTTAGAGTTTGGGAAAGGTACTAGGATCAGGTGCTTTT 1920  
QY 1921 GGAAAGTCAATGACCAAGCTTATGAAATTTAGCAAAACAGAGTCTCAATCCAGGT 1980  
DB 1921 GGAAAGTCAATGACCAAGCTTATGAAATTTAGCAAAACAGAGTCTCAATCCAGGT 1980  
QY 1981 GCGCTCAAAATGCTGAAAGAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAA 2040  
DB 1981 GCGCTCAAAATGCTGAAAGAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAA 2040  
QY 2041 CTCAGATGATGACCCAGCTGGAAAGCCAGAGAAATATGAGAACTGCTGGGGGGTGC 2100  
DB 2041 CTCAGATGATGACCCAGCTGGAAAGCCAGAGAAATATGAGAACTGCTGGGGGGTGC 2100  
QY 2101 ACATGTCAGGACCAATTTACTTATTTGAAATCTGTTGCTATGTTGATCTTCTCAAC 2160  
DB 2101 ACATGTCAGGACCAATTTACTTATTTGAAATCTGTTGCTATGTTGATCTTCTCAAC 2160  
QY 2161 TATCTAAGAGTAAAGAGAAATTTTCAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
DB 2161 TATCTAAGAGTAAAGAGAAATTTTCAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
QY 2221 AATTTCAAGTTTACCCCACTTTCCAAATCATCTCCAAATTCAGCATGCTGGTTCAAGA 2280  
DB 2221 AATTTCAAGTTTACCCCACTTTCCAAATCATCTCCAAATTCAGCATGCTGGTTCAAGA 2280  
QY 2281 GAAATTCAGATACACCCGGACTGGATCAATCTCAGGGCTTCATGGGAATTCATTTTAC 2340  
DB 2281 GAAATTCAGATACACCCGGACTGGATCAATCTCAGGGCTTCATGGGAATTCATTTTAC 2340

QY 2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAAAGGCTGGAAAGAGGAGGACTTTGAAT 2400  
DB 2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAAAGGCTGGAAAGAGGAGGACTTTGAAT 2400  
QY 2401 GTGCTTACATTTGAAGATCTTCTTTGCTTTGCTATCAAGTTGCCAAAGGAAATGGAATTT 2460  
DB 2401 GTGCTTACATTTGAAGATCTTCTTTGCTTTGCTATCAAGTTGCCAAAGGAAATGGAATTT 2460  
QY 2461 CTGGAATTTAAGTGTGTTTACAGAGACCTGGCCGCCAGGAACTGCTTCTCAGCCAC 2520  
DB 2461 CTGGAATTTAAGTGTGTTTACAGAGACCTGGCCGCCAGGAACTGCTTCTCAGCCAC 2520  
QY 2521 GGGAAAGTGGTGAAGATATGTGACTTTTGGATTGGCTCGAGATATCATGAGTGAATCCAA 2580  
DB 2521 GGGAAAGTGGTGAAGATATGTGACTTTTGGATTGGCTCGAGATATCATGAGTGAATCCAA 2580  
QY 2581 TATGTTGTGAGGGCAATGCCCGTCTGCTGTAAAATGATGGCCGCCGAAAGCCTGTTT 2640  
DB 2581 TATGTTGTGAGGGCAATGCCCGTCTGCTGTAAAATGATGGCCGCCGAAAGCCTGTTT 2640  
QY 2641 GAAGGCATCTACACCATTAAGAGTGTCTGCTCATATGGAATATTTACTGTGGGAAATC 2700  
DB 2641 GAAGGCATCTACACCATTAAGAGTGTCTGCTCATATGGAATATTTACTGTGGGAAATC 2700  
QY 2701 TTCTCACCTTGGTGTGAATCCTTACCTCGCATTCGGTGTGATGCTAACTTCTACAACTG 2760  
DB 2701 TTCTCACCTTGGTGTGAATCCTTACCTCGCATTCGGTGTGATGCTAACTTCTACAACTG 2760  
QY 2761 ATTTCAAAATGGAATTTAAAATGATCAGCCATTTATGTACAGAGAAATATACATTTATA 2820  
DB 2761 ATTTCAAAATGGAATTTAAAATGATCAGCCATTTTATGTACAGAGAAATATACATTTATA 2820  
QY 2821 ATGCAATCTGCTGGGCTTTGACTCAAGGAAACGGCCATCTTCCCTAAATTTGACTTGC 2880  
DB 2821 ATGCAATCTGCTGGGCTTTGACTCAAGGAAACGGCCATCTTCCCTAAATTTGACTTGC 2880  
QY 2881 TTTTGTAGTGTGCTGCGATGCAGAGAGAGGATGATCAGAAATGATGAGTGCCTG 2940  
DB 2881 TTTTGTAGTGTGCTGCGATGCAGAGAGAGGATGATCAGAAATGATGAGTGCCTG 2940  
QY 2941 GTTTCGGAATGCTTCAACCTTACCAAAACAGCGGACCTTTTCAGCAGAGAGATGATTTG 3000  
DB 2941 GTTTCGGAATGCTTCAACCTTACCAAAACAGCGGACCTTTTCAGCAGAGAGATGATTTG 3000  
QY 3001 GGGCTACTCTCTCCGAGGCTCAGGTGGAAGATTCGTAGAGGAACTTGTAGTTTAAAG 3060  
DB 3001 GGGCTACTCTCTCCGAGGCTCAGGTGGAAGATTCGTAGAGGAACTTGTAGTTTAAAG 3060  
QY 3061 ACTTCATCTCTCCCTATCCCTTAAAGGACTTTTGTAAAATCAAAATCATCTGTCAGAGCAG 3120  
DB 3061 ACTTCATCTCTCCCTATCCCTTAAAGGACTTTTGTAAAATCAAAATCATCTGTCAGAGCAG 3120  
QY 3121 CACTAAAAGAAATCTATTATCAACTGCTTCCAGCAGACTTTTCTAGAGAGCGCTCT 3180  
DB 3121 CACTAAAAGAAATCTATTATCAACTGCTTCCAGCAGACTTTTCTAGAGAGCGCTCT 3180  
QY 3181 GGGTTTACTCTTGTTTTCAAAGGACTTTTGTAAAATCAAAATCATCTGTCAGAGCAG 3240  
DB 3181 GGGTTTACTCTTGTTTTCAAAGGACTTTTGTAAAATCAAAATCATCTGTCAGAGCAG 3240  
QY 3241 GAGGAGCTGATTAATGAATTTTATTTGAGCAATGATCTGATCCAAAGGCTTCTCAGGCG 3300  
DB 3241 GAGGAGCTGATTAATGAATTTTATTTGAGCAATGATCTGATCCAAAGGCTTCTCAGGCG 3300  
QY 3301 GCTTGAAGTCAATTTGCTGCTGAGTACAGTATTTCTTGTAAATACATAAAACAAAGC 3360  
DB 3301 GCTTGAAGTCAATTTGCTGCTGAGTACAGTATTTCTTGTAAATACATAAAACAAAGC 3360  
QY 3361 ATTTTCTAAGGAGAGAGCTAAATGATTTTAAAGTCTATGTTTAAATTAATATGTA 3420  
DB 3361 ATTTTCTAAGGAGAGAGCTAAATGATTTTAAAGTCTATGTTTAAATTAATATGTA 3420  
QY 3421 TTTTTCAGCTATTTAGTGATATATTTTATGGGTTGGGAAATAAATTTCTACTACAGAAAAA 3480

Db 3421 TTTTTCAGCTATTAGTGATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAAA 3480  
QY 3481 AAAAAAAAAAAAAAAAAAAAAA 3501  
Db 3481 AAAAAAAAAAAAAAAAAAAAAA 3501

RESULT 3  
US-07-906-397A-3  
; Sequence 3, Application US/07906397A  
; Patent No. 5621090  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCLONE SYSTEMS INCORPORATED  
; STREET: 180 VARICK STREET  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/07/906,397A  
; FILING DATE: 19920626  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/813,593  
; FILING DATE: 24-DEC-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/793,065  
; FILING DATE: 15-NOV-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/728,913  
; FILING DATE: 28-JUN-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/679,666  
; FILING DATE: 02-APR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Feit, Irving N.  
; REGISTRATION NUMBER: 28,601  
; REFERENCE/DOCKET NUMBER: LEM-3-PPPPPP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-645-1405  
; TELEFAX: 212-645-2054  
; INFORMATION FOR SEQ ID NO. 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3501 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 58..3039  
US-07-906-397A-3

Query Match 100.0%; Score 3501; DB 1; Length 3501;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGAGGGGCGATCCGAGGGCTGGGCGCGCCCTGGGGGACCCCGGGCTCCGAGGCCCATG 60  
Db 1, CGAGGGGCGATCCGAGGGCTGGGCGCGCCCTGGGGGACCCCGGGCTCCGAGGCCCATG 60

QY 61 CCGGCGTTGGCGGCGACCGGGGACCGTGGCCCTGCTCGTTCGTTTTCTTCTGCAATGATA 120  
Db 61 CCGGCGTTGGCGGCGACCGGGGACCGTGGCCCTGCTCGTTCGTTTTCTTCTGCAATGATA 120  
QY 121 TTTGGGACTATTACAAATCAAGATCTGCCCTGTCATCAAGTGTCTTTTAAATCAATCATAAG 180  
Db 121 TTTGGGACTATTACAAATCAAGATCTGCCCTGTCATCAAGTGTCTTTTAAATCAATCATAAG 180  
QY 181 AACAAATGATTATCATCAGTGGGGAAGTCATCATATATCCCATGGTATCAGAAATCCCGGAA 240  
Db 181 AACAAATGATTATCATCAGTGGGGAAGTCATCATATATCCCATGGTATCAGAAATCCCGGAA 240  
QY 241 GACCTCGGTGTGCGTTGAGACCCGAGCTCAGGGACAGTGTACGAGCTGGCGTGTG 300  
Db 241 GACCTCGGTGTGCGTTGAGACCCGAGCTCAGGGACAGTGTACGAGCTGGCGTGTG 300  
QY 301 GAAGTGGATGTATCTGCTTCCATCACACTGCAAGTGTGCTGATGCCCGAGGGAACATT 360  
Db 301 GAAGTGGATGTATCTGCTTCCATCACACTGCAAGTGTGCTGATGCCCGAGGGAACATT 360  
QY 361 TCCTGTCTCTGGGTCTTTAAGCAAGCTCCCTGAAATGCCAGCCACATTTTGAATACAA 420  
Db 361 TCCTGTCTCTGGGTCTTTAAGCAAGCTCCCTGAAATGCCAGCCACATTTTGAATACAA 420  
QY 421 AACAGAGGAGTTGTTTCCATGTCATTTGAAATGACAGAAACCCCAAGCTGGAGATAC 480  
Db 421 AACAGAGGAGTTGTTTCCATGTCATTTGAAATGACAGAAACCCCAAGCTGGAGATAC 480  
QY 481 CTACTTTTATTCAGAGTGAAGCTACCAATTACAAATATTTTACAGTGAATGAAGA 540  
Db 481 CTACTTTTATTCAGAGTGAAGCTACCAATTACAAATATTTTACAGTGAATGAAGA 540  
QY 541 AATACCTCTGCTTTACACATTAAGAAGACCTTACTTTAGAAAAATGAAAAACCCAGAGGCC 600  
Db 541 AATACCTCTGCTTTACACATTAAGAAGACCTTACTTTAGAAAAATGAAAAACCCAGAGGCC 600  
QY 601 CTGGTCTGCATATCTCAGAGCGTTTCAGAGCCGATCGTGGAAATGGGTGCTTTGCGATCA 660  
Db 601 CTGGTCTGCATATCTCAGAGCGTTTCAGAGCCGATCGTGGAAATGGGTGCTTTGCGATCA 660  
QY 721 CATGAATTTATTCGGGACGGACATAAGGTGCTGTCAGAGAAATGAACTGGGCGAGGAATGC 780  
Db 721 CATGAATTTATTCGGGACGGACATAAGGTGCTGTCAGAGAAATGAACTGGGCGAGGAATGC 780  
QY 781 ACCAGGCTGTTCAATAGATCTTAAATCAAACTCCTCAGACCCACATTTGCCACATTTAT 840  
Db 781 ACCAGGCTGTTCAATAGATCTTAAATCAAACTCCTCAGACCCACATTTGCCACATTTAT 840  
QY 841 CTTAAAGTAGGGGAACCCCTTATGGATAAGTGCAAGCTGTTTATGTAACCATGGATTC 900  
Db 841 CTTAAAGTAGGGGAACCCCTTATGGATAAGTGCAAGCTGTTTATGTAACCATGGATTC 900  
QY 901 GGGCTCACCTGGGAATTAGAAAACAAAGCACTCGAGGAGGCAACTACTTTGAGATCAGT 960  
Db 901 GGGCTCACCTGGGAATTAGAAAACAAAGCACTCGAGGAGGCAACTACTTTGAGATCAGT 960  
QY 961 ACCTATTCAACAAACAGAACTATGATACGGATTTCTGTTTCTTTTGTATCATCAGTGGCA 1020  
Db 961 ACCTATTCAACAAACAGAACTATGATACGGATTTCTGTTTCTTTTGTATCATCAGTGGCA 1020  
QY 1021 AGAAACGACACCCGATACACTTGTTCCTTCAAAGCATCCAGTCAATCAGCTTTG 1080  
Db 1021 AGAAACGACACCCGATACACTTGTTCCTTCAAAGCATCCAGTCAATCAGCTTTG 1080  
QY 1081 GTTACCATCTAGGAAAGGATTTATAAATGCTACCAATTCAGTGAAGATTAATGAAT 1140  
Db 1081 GTTACCATCTAGGAAAGGATTTATAAATGCTACCAATTCAGTGAAGATTAATGAAT 1140

1141 GACCAATATGAAGAGTTTGTGTTTCTGTGTCAGGTTTAAAGCTACCCACAAATCAGATG 1200  
1141 GACCAATATGAAGAGTTTGTGTTTCTGTGTCAGGTTTAAAGCTACCCACAAATCAGATG 1200  
1201 ACCTGGACCTTCTCTGAAATCATTTTCTTGTGAGCAAAAGGCTTTGATAGCGATAC 1260  
1201 ACCTGGACCTTCTCTGAAATCATTTTCTTGTGAGCAAAAGGCTTTGATAGCGATAC 1260  
1261 AGCATATCCAAGTTTGTCAATCATTAAGCAACGCGAGGAGATATATTTCCATGCGAGAA 1320  
1261 AGCATATCCAAGTTTGTCAATCATTAAGCAACGCGAGGAGATATATTTCCATGCGAGAA 1320  
1321 AATGATGATGCCCAATTTACCAAAATGTTTCAAGCTGATATTAAGAAGGAACTCAAGTG 1380  
1321 AATGATGATGCCCAATTTACCAAAATGTTTCAAGCTGATATTAAGAAGGAACTCAAGTG 1380  
1381 CTGCGAAGCATCGGCAAGTCAAGCGTCTGTTTCTCGGATGATACCCATTTACCATCT 1440  
1381 CTGCGAAGCATCGGCAAGTCAAGCGTCTGTTTCTCGGATGATACCCATTTACCATCT 1440  
1441 TGGACCTGGAGAGTGTTCAGACAAGTCTCCAACTGCAAGAGAGATCACAGAGGA 1500  
1441 TGGACCTGGAGAGTGTTCAGACAAGTCTCCAACTGCAAGAGAGATCACAGAGGA 1500  
1501 GTCTGGAATGAAGAGTGTTCAGACAAGTGTTCGACAGTGGGTGTGAGCAGTACTCTA 1560  
1501 GTCTGGAATGAAGAGTGTTCAGACAAGTGTTCGACAGTGGGTGTGAGCAGTACTCTA 1560  
1561 AACATGATGAGCATTAAAGGGTCTCTGCTCAAGTGTGTCATACAAATTCCTCTGGC 1620  
1561 AACATGATGAGCATTAAAGGGTCTCTGCTCAAGTGTGTCATACAAATTCCTCTGGC 1620  
1621 ACATCTTGTGAGACGATCTTTAAACTCTCCAGGCCCTTCCCTTTCATCAAGACAAAC 1680  
1621 ACATCTTGTGAGACGATCTTTAAACTCTCCAGGCCCTTCCCTTTCATCAAGACAAAC 1680  
1681 ATCTCATTTATGCAACAAATTTGGTGTGTTGCTCTCTTCAATGCTGTTTAAACCTGTCTA 1740  
1681 ATCTCATTTATGCAACAAATTTGGTGTGTTGCTCTCTTCAATGCTGTTTAAACCTGTCTA 1740  
1741 ATTTGTGTCAGAGTACAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGTCAGAGTG 1800  
1741 ATTTGTGTCAGAGTACAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGTCAGAGTG 1800  
1801 ACCGGCTCTCAGATATGAGTACTTCTACGTTGATTTTCAAGAAATATGAATATGATCTC 1860  
1801 ACCGGCTCTCAGATATGAGTACTTCTACGTTGATTTTCAAGAAATATGAATATGATCTC 1860  
1861 AAATGGGAGTTTCCAAGAGAGAAAATTTAGAGTTTGGAAAGGTACTAGGATCAGGTGCTTTT 1920  
1861 AAATGGGAGTTTCCAAGAGAGAAAATTTAGAGTTTGGAAAGGTACTAGGATCAGGTGCTTTT 1920  
1921 GGAAGAGTATGAGCAAGCAAGCTTTATGGAATTAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
1921 GGAAGAGTATGAGCAAGCAAGCTTTATGGAATTAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
1981 GCCCTCAAAATGCTGAAGAGAGAAAACGACAGCTCTGAAAGAGAGGCACTCATGTCAAGAA 2040  
1981 GCCCTCAAAATGCTGAAGAGAGAAAACGACAGCTCTGAAAGAGAGGCACTCATGTCAAGAA 2040  
2041 CTCAGATGATGACCCAGCTGGGAAGCCACGAGAAATATTTGAAACCTGCTGGGGCGGTGC 2100  
2041 CTCAGATGATGACCCAGCTGGGAAGCCACGAGAAATATTTGAAACCTGCTGGGGCGGTGC 2100  
2101 ACATGTCAGGACCAATTTACTGATTTTGTGAATATGTTGATGCTGATGCTTCTCAAC 2160  
2101 ACATGTCAGGACCAATTTACTGATTTTGTGAATATGTTGATGCTGATGCTTCTCAAC 2160  
2161 TATCTAAGAGTAAAGAGAGAAAATTTACAGGACTTGGACAGAGATTTTCAAGGAAACAC 2220  
2161 TATCTAAGAGTAAAGAGAGAAAATTTACAGGACTTGGACAGAGATTTTCAAGGAAACAC 2220  
2221 AATTTTCAGTTTACCCTTTCCTCAATCACAATCCAAATTTCCAGCATGCTGTTCAAGA 2280

2221 AATTTTCAGTTTACCCTTTCCTCAATCACAATCCAAATTTCCAGCATGCTGTTCAAGA 2280  
2281 GAAGTTTCAGATACACCGGACTCGGATCAAAATCTCAGGGCTTCATGGAAATTCATTTTCAAC 2340  
2281 GAAGTTTCAGATACACCGGACTCGGATCAAAATCTCAGGGCTTCATGGAAATTCATTTTCAAC 2340  
2341 TCTGAAGATCAAAATTTGAATATGAAGAACCAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400  
2341 TCTGAAGATCAAAATTTGAATATGAAGAACCAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400  
2401 GTGCTTACATTTGAAGATCTTTTGTGTCATCAAGTTTGCACAAAGGAAATGGAATTTT 2460  
2401 GTGCTTACATTTGAAGATCTTTTGTGTCATCAAGTTTGCACAAAGGAAATGGAATTTT 2460  
2461 CTGGAATTTAAGTCTGTTGTTCAGAGAGCTTGGCCGCGCAGGAACTGCTTGTCAACCAC 2520  
2461 CTGGAATTTAAGTCTGTTGTTCAGAGAGCTTGGCCGCGCAGGAACTGCTTGTCAACCAC 2520  
2521 GGGAAAGTGTGAGATATGTGACTTTTGGATTTGGCTCGAGATATCATGAGTGAATTTCCAAAC 2580  
2521 GGGAAAGTGTGAGATATGTGACTTTTGGATTTGGCTCGAGATATCATGAGTGAATTTCCAAAC 2580  
2581 TATGTTCTCAGGGGCAATGCCCCGTCTGCTGTGTAATAATGGATGGCCCCCGAAAGCTGTTT 2640  
2581 TATGTTCTCAGGGGCAATGCCCCGTCTGCTGTGTAATAATGGATGGCCCCCGAAAGCTGTTT 2640  
2641 GAAGGCATCTACACCAATTAAGAGTGTCTGCTGTCATATGGAATATTTACTGTGGGAAATC 2700  
2641 GAAGGCATCTACACCAATTAAGAGTGTCTGCTGTCATATGGAATATTTACTGTGGGAAATC 2700  
2701 TTCTCATCTTGGTGTGAATCTTACCCTGGCAATCCGGTTGATGCTTCTTCAAAAATG 2760  
2701 TTCTCATCTTGGTGTGAATCTTACCCTGGCAATCCGGTTGATGCTTCTTCAAAAATG 2760  
2761 ATTCAAAATGGATTTAAATATGGATCAGCCATTTTATGCTACAGAGAAATATACATTTATA 2820  
2761 ATTCAAAATGGATTTAAATATGGATCAGCCATTTTATGCTACAGAGAAATATACATTTATA 2820  
2821 ATGCAATCTCTGCTGGCTTTTGAATCAAGGAAACGGCCATCTCTTCCCTTAAATTTGACTCG 2880  
2821 ATGCAATCTCTGCTGGCTTTTGAATCAAGGAAACGGCCATCTCTTCCCTTAAATTTGACTCG 2880  
2881 TTTTTCAGATGTCAGTGGCAGATGTCAGAGAGCGATGATCAGAAATGTCGATGTCGCTG 2940  
2881 TTTTTCAGATGTCAGTGGCAGATGTCAGAGAGCGATGATCAGAAATGTCGATGTCGCTG 2940  
2941 GTTTTCGGAATGTCCTCAACCTTACCAGGAGCGCTTTCAGAGAGAGATGATGATTTG 3000  
2941 GTTTTCGGAATGTCCTCAACCTTACCAGGAGCGCTTTCAGAGAGAGATGATGATTTG 3000  
3001 GGGCTACTCTCTCCGCGAGCTCAGGTCTGAGAGATTCGTAGAGGAAACAAATTTAGTTTAAAG 3060  
3001 GGGCTACTCTCTCCGCGAGCTCAGGTCTGAGAGATTCGTAGAGGAAACAAATTTAGTTTAAAG 3060  
3061 ACTTCATCCCTCCACTATCCCTAACAGGCTGTAGATTTACCAAAACAAGATTTAAATTTTCA 3120  
3061 ACTTCATCCCTCCACTATCCCTAACAGGCTGTAGATTTACCAAAACAAGATTTAAATTTTCA 3120  
3121 CACTTAAAGAGAAATCTATTTATCACTGCTTCCAGAGCTTTTCTCTAGAGGCGCTCT 3180  
3121 CACTTAAAGAGAAATCTATTTATCACTGCTTCCAGAGCTTTTCTCTAGAGGCGCTCT 3180  
3181 GCGTTTACTCTTGTTTTCAAAGGGACTTTTGTAAATCAAAATCATCTCTGTCAAGGCGAG 3240  
3181 GCGTTTACTCTTGTTTTCAAAGGGACTTTTGTAAATCAAAATCATCTCTGTCAAGGCGAG 3240  
3241 GAGGAGCTGATAATGAACCTTTATTTGGAGCATTTGATCTGCAATCCAGGCTTCTCAGGCGG 3300  
3241 GAGGAGCTGATAATGAACCTTTATTTGGAGCATTTGATCTGCAATCCAGGCTTCTCAGGCGG 3300  
3301 GCTTGAAGTGAATTTGTTGTTACCTGGAAGTACAGTATATTTCTTGTAAATACATAAAACAAAGC 3360

Db 3301 GCTTGAAGTAACTGTACCTGAAGTACAGTATATTTCTGTAAATACATAAAACAAAGC 3360  
Qy 3361 ATTTTCTAAGAGAGCTAATATGATTTTAACTCTATGTTTAAATAATATGTAAA 3420  
Db 3361 ATTTTCTAAGAGAGCTAATATGATTTTAACTCTATGTTTAAATAATATGTAAA 3420  
Qy 3421 TTTTTCAGCTATTAGTCATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAA 3480  
Db 3421 TTTTTCAGCTATTAGTCATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAA 3480  
Qy 3481 AAAAAAAAAAAAAAAAAAAAAA 3501  
Db 3481 AAAAAAAAAAAAAAAAAAAAAA 3501

## RESULT 4

US-08-601-891-3  
; Sequence 3, Application US/08601891  
; Patent No. 5747651  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ImClone Systems Incorporated  
; STREET: 180 Varick Street  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/601,891  
; FILING DATE: 15-FEB-1996  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/977,451  
; FILING DATE: 19-NOV-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/906,397  
; FILING DATE: 26-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US PCT/US92/05401  
; FILING DATE: 26-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: TW 81102961  
; FILING DATE: 15-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US PCT/US92/02750  
; FILING DATE: 02-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/813,593  
; FILING DATE: 24-DEC-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/793,665  
; FILING DATE: 15-NOV-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/728,913  
; FILING DATE: 28-JUN-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/679,666  
; FILING DATE: 02-APR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Feit, Irving N.  
; REGISTRATION NUMBER: 28,601  
; REFERENCE/DOCKET NUMBER: LEM-3-7P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-645-1405

TELEFAX: 212-645-2054  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3501 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 58..3039  
; FEATURE:  
; NAME/KEY: mat\_peptide  
; LOCATION: 139..3036  
; FEATURE:  
; NAME/KEY: sig\_peptide  
; LOCATION: 58..138  
US-08-601-891-3  
Query Match 100.0%; Score 3501; DB 1; Length 3501;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CGAGGCGGATCCGAGGCGTGGGCGGCGCCCTGGGGACCCCGGGCTCCGAGAGCCCATG 60  
Db 1 CGAGGCGGATCCGAGGCGTGGGCGGCGCCCTGGGGACCCCGGGCTCCGAGAGCCCATG 60  
Qy 61 CGGCGGTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120  
Db 61 CGGCGGTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120  
Qy 121 TTTGGGACTATTACAAATCAAGATCTGCTCTGTGATCAAGTGTGTTTAAATCAATCAATG 180  
Db 121 TTTGGGACTATTACAAATCAAGATCTGCTCTGTGATCAAGTGTGTTTAAATCAATCAATG 180  
Qy 181 AACAAATGATTCAGTGGGGAAGTCATCATATATCCATGATCAAGATCCCGCGAA 240  
Db 181 AACAAATGATTCAGTGGGGAAGTCATCATATATCCATGATCAAGATCCCGCGAA 240  
Qy 241 GACCTCGGTTGGTGGAGACCCAGAGCTCAGGACAGTGTACGAAGCTGCCCTGTG 300  
Db 241 GACCTCGGTTGGTGGAGACCCAGAGCTCAGGACAGTGTACGAAGCTGCCCTGTG 300  
Qy 301 GAAGTGGATGTATCTGCTTCCATCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
Db 301 GAAGTGGATGTATCTGCTTCCATCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
Qy 361 TCCTGTCTCTGGGCTTTTAAAGCAGAGCTCCCTGAAATTCGCCAGCCACATTTTGTATTTACAA 420  
Db 361 TCCTGTCTCTGGGCTTTTAAAGCAGAGCTCCCTGAAATTCGCCAGCCACATTTTGTATTTACAA 420  
Qy 421 AACAGAGAGTGTGTTCCATGCTGTTTGAATAATGACAGAAACCCAGAGCTGGAGAAATAC 480  
Db 421 AACAGAGAGTGTGTTCCATGCTGTTTGAATAATGACAGAAACCCAGAGCTGGAGAAATAC 480  
Qy 481 CTACTTTTATTCAGAGTGAAGTACCAATTTACACAAATTTGTTTACAGTGAATGATA 540  
Db 481 CTACTTTTATTCAGAGTGAAGTACCAATTTACACAAATTTGTTTACAGTGAATGATA 540  
Qy 541 AATACCCCTGCTTACACATTAAGAGAGCTTTACTTTTGAAGAAATGAGAAACCCAGAGCGCC 600  
Db 541 AATACCCCTGCTTACACATTAAGAGAGCTTTACTTTTGAAGAAATGAGAAACCCAGAGCGCC 600  
Qy 601 CTGGTCTGCATATCTGAGAGCGTTCCAGAGCGGATCGTGGAAATGGGTGCTTTCGCGATTCA 660  
Db 601 CTGGTCTGCATATCTGAGAGCGTTCCAGAGCGGATCGTGGAAATGGGTGCTTTCGCGATTCA 660  
Qy 661 CAGGGGGAAGCTGTAAAGAGAAAGTCCAGCTGTGTTTAAAGAGGAGGAAAAAGTGCTT 720  
Db 661 CAGGGGGAAGCTGTAAAGAGAAAGTCCAGCTGTGTTTAAAGAGGAGGAAAAAGTGCTT 720



2881	Db		TTTTTAGGATGT	CAGCTGGCAGATG	GCAGAGAAAGCGATGTAT	CAGAAATGGGATGCCCGT	2940
2941	Qy		GTTTTCGGAATGT	CCTCACA	CCTACCAAAA	CAGGCGACCTTTTCAGCAGAGATGGGATTTG	3000
2941	Db		GTTTTCGGAATGT	CCTCACA	CCTACCAAAA	CAGGCGACCTTTTCAGCAGAGATGGGATTTG	3000
3001	Qy		GGGCTACTCTCT	CGCGAGGCTC	AGAGTTCGAAGATTCGTAGAGGAACAATTTAGTTTAAAGG	3060	
3001	Db		GGGCTACTCTCT	CGCGAGGCTC	AGAGTTCGAAGATTCGTAGAGGAACAATTTAGTTTAAAGG	3060	
3061	Qy		ACTTTCATCCCT	CCACCTATCCCT	TAAACAGGCTGTAGATTACCAAAA	CAAGATTAATTTTCAT	3120
3061	Db		ACTTTCATCCCT	CCACCTATCCCT	TAAACAGGCTGTAGATTACCAAAA	CAAGATTAATTTTCAT	3120
3121	Qy		CACATAAAGAAAA	TCTATTATCA	CTGCTGCTTCA	CCAGACTTTTCTCTAGAGCCGCTCT	3180
3121	Db		CACATAAAGAAAA	TCTATTATCA	CTGCTGCTTCA	CCAGACTTTTCTCTAGAGCCGCTCT	3180
3181	Qy		CGGTTTACTCT	TGTTTTTCAA	AGGGACTTTTGTAAAAATCAAATCA	TCTGTCACAAAGGCAG	3240
3181	Db		CGGTTTACTCT	TGTTTTTCAA	AGGGACTTTTGTAAAAATCAAATCA	TCTGTCACAAAGGCAG	3240
3241	Qy		GAGGAGCTGATA	TATGA	ACTTTATTTGAGAGCAT	TGATCTGCATCCAGGCCCTTCTCAGGCCG	3300
3241	Db		GAGGAGCTGATA	TATGA	ACTTTATTTGAGAGCAT	TGATCTGCATCCAGGCCCTTCTCAGGCCG	3300
3301	Qy		GCTTTGAGTGAAT	CTGTGTACCT	TGAAGTACAGTATATCTCTGTAAATACATAAAA	CAAAAAGC	3360
3301	Db		GCTTTGAGTGAAT	CTGTGTACCT	TGAAGTACAGTATATCTCTGTAAATACATAAAA	CAAAAAGC	3360
3361	Qy		ATTTTGTCTAAGG	AGAGCTAA	TGATTTTTTTTAA	GTCTATGTTTTTAAAAATAATATGTAAA	3420
3361	Db		ATTTTGTCTAAGG	AGAGCTAA	TGATTTTTTTTAA	GTCTATGTTTTTAAAAATAATATGTAAA	3420
3421	Qy		TTTTTTCAGCTAT	TTTTAGTGATA	TATTTTTTATGGTGGGAATAAAA	TTTCTCTACTACAGAAAAA	3480
3421	Db		TTTTTTCAGCTAT	TTTTAGTGATA	TATTTTTTATGGTGGGAATAAAA	TTTCTCTACTACAGAAAAA	3480
3481	Qy		AAAAA	AAAAAAAAAAAAAAAAA	AAAAA	3501	
3481	Db		AAAAA	AAAAAAAAAAAAAAAAA	AAAAA	3501	

RESULT 5  
US-09-021-324-3  
; Sequence 3, Application US/09021324  
; Patent No. 5912133  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ImClone Systems Incorporated  
; STREET: 180 Varick Street  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/021,324  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/977,451  
; FILING DATE: 1992-11-19  
; PRIOR APPLICATION DATA:

```

APPLICATION NUMBER: US 07/906,397
FILING DATE: 26-JUN-1992
PRIORITY: 26-JUN-1992
APPLICATION DATA:
APPLICATION NUMBER: US PCT/US92/05401
FILING DATE: 26-JUN-1992
PRIORITY: 26-JUN-1992
APPLICATION DATA:
APPLICATION NUMBER: TW 81102961
FILING DATE: 15-APR-1992
PRIORITY: 15-APR-1992
APPLICATION DATA:
APPLICATION NUMBER: US PCT/US92/02750
FILING DATE: 02-APR-1992
PRIORITY: 02-APR-1992
APPLICATION DATA:
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
PRIORITY: 24-DEC-1991
APPLICATION DATA:
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
PRIORITY: 15-NOV-1991
APPLICATION DATA:
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
PRIORITY: 28-JUN-1991
APPLICATION DATA:
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
PRIORITY: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Peit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3501 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 58..3039
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 139..3036
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..138
US-09-021-324-3

Query Match 100.0%; Score 3501; DB 2; Length 3501;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGAGGCGGCATCCGAGGCGTGGGCCGCGCCCTGGGGAGACCCCGGGCTCCGAGGCCATG 60
Db 1 CGAGGCGGCATCCGAGGCGTGGGCCGCGCCCTGGGGAGACCCCGGGCTCCGAGGCCATG 60
Qy 61 CCGCGTTGGGCGCGAGCGGGGACCGTCCGCTGCTGCTGTTTTTCTGCAATGATA 120
Db 61 CCGCGTTGGGCGCGAGCGGGGACCGTCCGCTGCTGCTGTTTTTCTGCAATGATA 120
Qy 121 TTTCGGGACTATTACAAATCAAGATCTCGCTGTGATCAAGTGTGTTTTTAATCAATCAATG 180
Db 121 TTTCGGGACTATTACAAATCAAGATCTCGCTGTGATCAAGTGTGTTTTTAATCAATCAATG 180
Qy 181 AACAAATGATTTCATCAGTGGGGAGTTCATCATCATATCCCATGGTATCAGAAATCCCGGAA 240
Db 181 AACAAATGATTTCATCAGTGGGGAGTTCATCATCATATCCCATGGTATCAGAAATCCCGGAA 240
Qy 241 GACCTCGGGTGTGCTTGAGACCCCAAGCTCAGGACAGTGTACGAAGCTGCGCTGTG 300

```



Db 241 GACCTCGGCTGTGCTTTGAGACCCAGAGCTCAGGCA CAGTGTACGAAGCTCGCGCTGTG 300  
Qy 301 GAAGTGAATCTATCTGCTTCCATCACACTGCAAGTGTGTCGATGCCAGGGAACATT 360  
Db 301 GAAGTGAATGTATCTGCTTCCATCACACTGCAAGTGTGTCGATGCCAGGGAACATT 360  
Qy 361 TCCTGTCTCTGGGTCTTTAAAGCAGAGCTCCCTGAAATGCCAGGCACACATTTGATTTACAA 420  
Db 361 TCCTGTCTCTGGGTCTTTAAAGCAGAGCTCCCTGAAATGCCAGGCACACATTTGATTTACAA 420  
Qy 421 AACAGAGGAGTGTTCATCGTGTCAATTTGAAATGACAGAAACCCAAAGCTGGAGATAC 480  
Db 421 AACAGAGGAGTGTTCATCGTGTCAATTTGAAATGACAGAAACCCAAAGCTGGAGATAC 480  
Qy 481 CTACTTTTATTCAGAGTGAAGCTTACCATAATTTACACAATATTTGTTTACAGTGAATAGA 540  
Db 481 CTACTTTTATTCAGAGTGAAGCTTACCATAATTTACACAATATTTGTTTACAGTGAATAGA 540  
Qy 541 AATACCCTGTCTTACACATTTAAGAGACCTTACTTTAGAAAATGGAACCCAGAGGCC 600  
Db 541 AATACCCTGTCTTACACATTTAAGAGACCTTACTTTAGAAAATGGAACCCAGAGGCC 600  
Qy 601 CTGGTCTGCATATCTGAGAGCTTCCAGAGCGATCGTGAATGGGTGCTTTGCGATTC 660  
Db 601 CTGGTCTGCATATCTGAGAGCGTTCAGAGCGATCGTGAATGGGTGCTTTGCGATTC 660  
Qy 661 CAGGGGAAAGCTGTAAGAGAAAGTCCAGCTGTTGTTAAAGAGGAGAAAAGTCTT 720  
Db 661 CAGGGGAAAGCTGTAAGAGAAAGTCCAGCTGTTGTTAAAGAGGAGAAAAGTCTT 720  
Qy 721 CATGAATATTTGGGACGGACATGAAGTGTGTCGCCAGAAATGAATGAGGAGGAAATGC 780  
Db 721 CATGAATATTTGGGACGGACATGAAGTGTGTCGCCAGAAATGAATGAGGAGGAAATGC 780  
Qy 781 ACCAGGCTGTTTCAATATAGATCTAAATCAAACTCCTCAGACCACATTTGCCACAAATATT 840  
Db 781 ACCAGGCTGTTTCAATATAGATCTAAATCAAACTCCTCAGACCACATTTGCCACAAATATT 840  
Qy 841 CTTAAAGTAGGGAAACCCCTTATGATGAAGAGTGTGTAAGCTGTTGATGTAACCATGGATTC 900  
Db 841 CTTAAAGTAGGGAAACCCCTTATGATGAAGAGTGTGTAAGCTGTTGATGTAACCATGGATTC 900  
Qy 901 GGGCTCACCTGGGAATTTAGAAAACAAAGCACTCAGGAGGGCAACTACTTTGAGATGAGT 960  
Db 901 GGGCTCACCTGGGAATTTAGAAAACAAAGCACTCAGGAGGGCAACTACTTTGAGATGAGT 960  
Qy 961 ACCTATTCAACAAACAGAACTATGATACGGATCTGTTTGTGTTGTTGATCATCAGTGGCA 1020  
Db 961 ACCTATTCAACAAACAGAACTATGATACGGATCTGTTTGTGTTGTTGATCATCAGTGGCA 1020  
Qy 1021 AGAAACGACACCGGATACACTGTTGCTCTTTCAAAGCATCCAGTCAATCAGCTTTG 1080  
Db 1021 AGAAACGACACCGGATACACTGTTGCTCTTTCAAAGCATCCAGTCAATCAGCTTTG 1080  
Qy 1081 GTTACCATCTAGGAAAGGATTTATAATGCTACCAATTCAGTGAAGATTATGAAT 1140  
Db 1081 GTTACCATCTAGGAAAGGATTTATAATGCTACCAATTCAGTGAAGATTATGAAT 1140  
Qy 1141 GACCAATATCAAGAGTTTGTGTTTCTGTCAGGTTTAAAGCTTACCACAAATCAGATGT 1200  
Db 1141 GACCAATATCAAGAGTTTGTGTTTCTGTCAGGTTTAAAGCTTACCACAAATCAGATGT 1200  
Qy 1201 ACSTGGACCTTCTCTGAAATATCTTTCTGTGAGCAAAAGGCTTTGATAACGGATAC 1260  
Db 1201 ACSTGGACCTTCTCTGAAATATCTTTCTGTGAGCAAAAGGCTTTGATAACGGATAC 1260  
Qy 1261 AGCATATCCAAGTTTTCGAATCATAGCAACAGCCAGGAGAAATATATTTCCATGCGAGAA 1320  
Db 1261 AGCATATCCAAGTTTTCGAATCATAGCAACAGCCAGGAGAAATATATTTCCATGCGAGAA 1320  
Qy 1321 AATGATGATGCCCAATTTTACCAAAATGTTTCAACGCTGAAATATAAGAGGAAACCTCAAGT 1380  
Db 1321 AATGATGATGCCCAATTTTACCAAAATGTTTCAACGCTGAAATATAAGAGGAAACCTCAAGT 1380

Qy 1381 CTCGCAAGACATCGCAAGTCAGGCGTCTCTGTTTCTCGATGATACCCATTACCATCT 1440  
Db 1381 CTCGCAAGACATCGCAAGTCAGGCGTCTCTGTTTCTCGATGATACCCATTACCATCT 1440  
Qy 1441 TGAACCTGGAAGAGTGTTCAGACAGTCTCCCAACTGCCACAGAGATCACAAGAGA 1500  
Db 1441 TGAACCTGGAAGAGTGTTCAGACAGTCTCCCAACTGCCACAGAGATCACAAGAGA 1500  
Qy 1501 GTCTGGNATGAAGGCTTACAGAAAAGTGTGGAAGTGGGTGTCGAGCAGTACTCTA 1560  
Db 1501 GTCTGGNATGAAGGCTTACAGAAAAGTGTGGAAGTGGGTGTCGAGCAGTACTCTA 1560  
Qy 1561 AACATGAGTCAAGCATAAAAGGCTTCCCTGCTCAAGTGTGTCATACAAATCCCTTGGC 1620  
Db 1561 AACATGAGTCAAGCATAAAAGGCTTCCCTGCTCAAGTGTGTCATACAAATCCCTTGGC 1620  
Qy 1621 ACATCTTGTGAGACGATCCCTTTAAACTCTCCAGGCCCTTCCCTTATCCAGACAAC 1680  
Db 1621 ACATCTTGTGAGACGATCCCTTTAAACTCTCCAGGCCCTTCCCTTATCCAGACAAC 1680  
Qy 1681 ATCTCATTTCTATGCAACAATTTGGTGTGTTGCTCTCTTCAATGTCGTTTAAACCTGCTA 1740  
Db 1681 ATCTCATTTCTATGCAACAATTTGGTGTGTTGCTCTCTTCAATGTCGTTTAAACCTGCTA 1740  
Qy 1741 ATTTGTCAACAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGCTACAGTG 1800  
Db 1741 ATTTGTCAACAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGCTACAGTG 1800  
Qy 1801 ACCGGCTCTCAGATAATGAGTACTTCTAGCTGATTTTCAGAGAAATATGAATATGATCTC 1860  
Db 1801 ACCGGCTCTCAGATAATGAGTACTTCTAGCTGATTTTCAGAGAAATATGAATATGATCTC 1860  
Qy 1861 AATGGAGTTCACAGAGAAAATTTAGAGTTTGGGAAAGTACTAGGATCAGGTGCTTTT 1920  
Db 1861 AATGGAGTTCACAGAGAAAATTTAGAGTTTGGGAAAGTACTAGGATCAGGTGCTTTT 1920  
Qy 1921 GGAAGAGTGAAGCAACAGCTTATGGAATAGCAAAACAGAGTCTCAATCCAGTT 1980  
Db 1921 GGAAGAGTGAAGCAACAGCTTATGGAATAGCAAAACAGAGTCTCAATCCAGTT 1980  
Qy 1981 GCCTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGGCACTCATGTCAAG 2040  
Db 1981 GCCTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGGCACTCATGTCAAG 2040  
Qy 2041 CTCAGATGATGACCCAGCTGGGAGCCACGAGAAATTTGTGAACCTGCTGGGGCGTGC 2100  
Db 2041 CTCAGATGATGACCCAGCTGGGAGCCACGAGAAATTTGTGAACCTGCTGGGGCGTGC 2100  
Qy 2101 ACACTGTCAAGACCAATTTACTGATTTTGAATACTGTTGTGCTATGTTGATCTTCTAAC 2160  
Db 2101 ACACTGTCAAGACCAATTTACTGATTTTGAATACTGTTGTGCTATGTTGATCTTCTAAC 2160  
Qy 2161 TATCTAAGAGTAAAGAGAAAATTTTCAGAGACTTTGGACAGAGATTTTCAAGAAAC 2220  
Db 2161 TATCTAAGAGTAAAGAGAAAATTTTCAGAGACTTTGGACAGAGATTTTCAAGAAAC 2220  
Qy 2221 AATTTCAAGTAAAGAGAAAATTTTCCAAATCACATCCAAATTCAGATGCTGTTCAAGA 2280  
Db 2221 AATTTCAAGTAAAGAGAAAATTTTCCAAATCACATCCAAATTCAGATGCTGTTCAAGA 2280  
Qy 2281 GAAGTTCAGATACACCGGACTCGGATCAATCTCAGGGTTCATGGAAATTCATTTCAAC 2340  
Db 2281 GAAGTTCAGATACACCGGACTCGGATCAATCTCAGGGTTCATGGAAATTCATTTCAAC 2340  
Qy 2341 TCTGAGATCAAAATGGAATGAAAACCAAAAGGCTGGAGAGGAGGAGGACTTGAAT 2400  
Db 2341 TCTGAGATCAAAATGGAATGAAAACCAAAAGGCTGGAGAGGAGGAGGACTTGAAT 2400  
Qy 2401 GTGCTTACATTTGAAGATCTTTTGTGTTGCTATCATCAAGTTGCCAAAGGAATGGAAAT 2460  
Db 2401 GTGCTTACATTTGAAGATCTTTTGTGTTGCTATCATCAAGTTGCCAAAGGAATGGAAAT 2460



2461 CTGGAATTTAAGTCTGTGTTTCAGAGACCTGGCGCCAGGACGCTGTGTCAACCCAC 2520  
 Db CTGGAATTTAAGTCTGTGTTTCAGAGACCTGGCGCCAGGACGCTGTGTCAACCCAC 2520  
 Qy GGGAAAGTGGTGAAGATATGATGACTTTGGATGGCTCGAGATATCATGAGTATCCAAAC 2580  
 Db GGGAAAGTGGTGAAGATATGATGACTTTGGATGGCTCGAGATATCATGAGTATCCAAAC 2580  
 Qy TATGTTCTCAGGGGCAATGCCCCGTCTGCTGTAAATGGATGGCCCCGAAAGCCTGTTT 2640  
 Db TATGTTCTCAGGGGCAATGCCCCGTCTGCTGTAAATGGATGGCCCCGAAAGCCTGTTT 2640  
 Qy GAAGGCATCTACACCATTAAGAGTGTCTGGTCTATGGAATATTAATCTGTGGAAATC 2700  
 Db GAAGGCATCTACACCATTAAGAGTGTCTGGTCTATGGAATATTAATCTGTGGAAATC 2700  
 Qy TTCTCACTGTGTGTAATCCTTACCCTGGCAATCCGGTTGATGCTAACTTCTACAAACTG 2760  
 Db TTCTCACTGTGTGTAATCCTTACCCTGGCAATCCGGTTGATGCTAACTTCTACAAACTG 2760  
 Qy ATTCAAAATGGATTTAAATGGATGAGCCATTTTATGCTCAGAGAAATATACATTATA 2820  
 Db ATTCAAAATGGATTTAAATGGATGAGCCATTTTATGCTCAGAGAAATATACATTATA 2820  
 Qy ATGCAATCTCTGGGCTTTTGAATCAAGGAAACGGCCATCCTTCCCTAAATTTGACTTCG 2880  
 Db ATGCAATCTCTGGGCTTTTGAATCAAGGAAACGGCCATCCTTCCCTAAATTTGACTTCG 2880  
 Qy TTTTATGAGTGTGAGTGGAGATGTCAGAGAGCGATGATGATGATGATGATGATGATG 2940  
 Db TTTTATGAGTGTGAGTGGAGATGTCAGAGAGCGATGATGATGATGATGATGATGATG 2940  
 Qy GTTTCGGAATCTCTCACACCTACCAAAACAGGCGACCTTTTTCAGCAGAGAGATGGATTG 3000  
 Db GTTTCGGAATCTCTCACACCTACCAAAACAGGCGACCTTTTTCAGCAGAGAGATGGATTG 3000  
 Qy GGGCTACTCTCTCGCGAGGCTCAGTCTGAAGATGCTGAGAGAACAAATTTAGTTTAAAG 3060  
 Db GGGCTACTCTCTCGCGAGGCTCAGTCTGAAGATGCTGAGAGAACAAATTTAGTTTAAAG 3060  
 Qy ACTTCACTCCCTCCACTATCCTTAACAGCTGTAGATTAACAAACAGAGATTAAATTCAT 3120  
 Db ACTTCACTCCCTCCACTATCCTTAACAGCTGTAGATTAACAAACAGAGATTAAATTCAT 3120  
 Qy CACTAAAGAAATCTATTATCACTGCTTCAACAGACTTTTCTCTAGAGCGGCTCT 3180  
 Db CACTAAAGAAATCTATTATCACTGCTTCAACAGACTTTTCTCTAGAGCGGCTCT 3180  
 Qy GCGTTTACTCTGTTTCAAAAGGACTTTTGTAAATCAAAATCAATCTCTGTCACAGGCG 3240  
 Db GCGTTTACTCTGTTTCAAAAGGACTTTTGTAAATCAAAATCAATCTCTGTCACAGGCG 3240  
 Qy GAGAGCTGATTAAGAACTTTTATGGAGCATTTGATCTGCATCCAGGCTTCTCAGGCGG 3300  
 Db GAGAGCTGATTAAGAACTTTTATGGAGCATTTGATCTGCATCCAGGCTTCTCAGGCGG 3300  
 Qy GCTTGAGTGAATGTTGTTACCTGAACTACAGTATATCTTGTAAATACATAAAACAAAGC 3360  
 Db GCTTGAGTGAATGTTGTTACCTGAACTACAGTATATCTTGTAAATACATAAAACAAAGC 3360  
 Qy ATTTTGTGAAGAGAGCTAATATGATTTTTTAACTCTATGTTTTTAAATATATGTAATA 3420  
 Db ATTTTGTGAAGAGAGCTAATATGATTTTTTAACTCTATGTTTTTAAATATATGTAATA 3420  
 Qy TTTTTCAGCTATTAGTATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAAA 3480  
 Db TTTTTCAGCTATTAGTATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAAA 3480  
 Qy AAAAAAAAAAAAAAAAAAAAAA 3501  
 Db AAAAAAAAAAAAAAAAAAAAAA 3501

RESULT 6  
 PCT-US92-05401-3  
 ; Sequence 3, Application PC/TUS9205401  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lemischka, Ihor R.  
 ; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
 ; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: INCLONE SYSTEMS INCORPORATED  
 ; STREET: 180 VARICK STREET  
 ; CITY: NEW YORK  
 ; STATE: NEW YORK  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10014  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US92/05401  
 ; FILING DATE: 19920626  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Felt, Irving N.  
 ; REGISTRATION NUMBER: 28,601  
 ; REFERENCE/DOCKET NUMBER: LEM-3-PPPPPT  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 212-645-1405  
 ; TELEFAX: 212-645-2054  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 3501 base pairs  
 ; TYPE: NUCLEIC ACID  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; HYPOTHETICAL: NO  
 ; ANTI-SENSE: NO  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: join(1..3036, 3040..3111, 3115..3123, 3127..3168,  
 ; LOCATION: 3172..3303, 3307..3348, 3352..3378, 3382..3501)  
 PCT-US92-05401-3  
 Query Match 100.0%; Score 3501; DB 5; Length 3501;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 CGAGGCGGCATCCGAGGCTGGGCGCGCCCTGGGGACCCCGGCTCCGAGGCCATG 60  
 Db 1 CGAGGCGGCATCCGAGGCTGGGCGCGCCCTGGGGACCCCGGCTCCGAGGCCATG 60  
 Qy 61 CCGGCGTTGGCGCGAGCGGGGACCGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
 Db 61 CCGGCGTTGGCGCGAGCGGGGACCGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
 Qy 121 TTTGGGACTATTACAAATCAAGATCTGCTGCTGATCAAGTGTGTTTAAATCAATCAATG 180  
 Db 121 TTTGGGACTATTACAAATCAAGATCTGCTGCTGATCAAGTGTGTTTAAATCAATCAATG 180  
 Qy 181 AACATGATTCATCAGTGGGAAAGTCATCATCATATCCATGATCAGAAATCCCGGAA 240  
 Db 181 AACATGATTCATCAGTGGGAAAGTCATCATCATATCCATGATCAGAAATCCCGGAA 240  
 Qy 241 GACCTCGGCTGCTGTTGAGACCCAGAGCTCAGGAGCAGTGTACGAAAGTCCCGCTGTG 300  
 Db 241 GACCTCGGCTGCTGTTGAGACCCAGAGCTCAGGAGCAGTGTACGAAAGTCCCGCTGTG 300  
 Qy 301 GAAGTGGATGTATCTGCTTCCATCAGACTGCAAGTGTGCTGATGCCCGGGAACATT 360  
 Db 301 GAAGTGGATGTATCTGCTTCCATCAGACTGCAAGTGTGCTGATGCCCGGGAACATT 360

QY 361 TCCTGTCTCTGGGCTTTTAAACACAGCTCCCTGAATTCGCCAGCCACATTTTGGATTACAA 420  
DB 361 TCCTGTCTCTGGGCTTTTAAACACAGCTCCCTGAATTCGCCAGCCACATTTTGGATTACAA 420  
QY 421 AACAGAGAGTGTTCATGCTTCAATTTTGAATTTGAGAAATGACAGAAACCAAGCTGGAGATAC 480  
DB 421 AACAGAGAGTGTTCATGCTTCAATTTTGAATTTGAGAAATGACAGAAACCAAGCTGGAGATAC 480  
QY 481 CTACTTTTATTCAGAGTGAAGCTACCAATTTACACAATTTGTGTACAGTGAGTATAGA 540  
DB 481 CTACTTTTATTCAGAGTGAAGCTACCAATTTACACAATTTGTGTACAGTGAGTATAGA 540  
QY 541 AATACCCCTGCTTTACACATTAAGAGAGCTTTACTTTTGAATAAATGGAATAACCCAGAGCC 600  
DB 541 AATACCCCTGCTTTACACATTAAGAGAGCTTTACTTTTGAATAAATGGAATAACCCAGAGCC 600  
QY 601 CTGCTCTGCATATCTGAGAGCGTTCCAGAGCGGATCGTGGAAATGGGTCTTTGGATTCA 660  
DB 601 CTGCTCTGCATATCTGAGAGCGTTCCAGAGCGGATCGTGGAAATGGGTCTTTGGATTCA 660  
QY 661 CAGGGGAAAGCTGTAAGAAAGAAAGTCCAGCTGTTTAAAGAGGAGGAAAAAGTGCTT 720  
DB 661 CAGGGGAAAGCTGTAAGAAAGAAAGTCCAGCTGTTTAAAGAGGAGGAAAAAGTGCTT 720  
QY 721 CATGAAATTTTGGGAGCGGACATAGGTGCTGCCAGAAATGAATGAACTGGGAGGAAATGC 780  
DB 721 CATGAAATTTTGGGAGCGGACATAGGTGCTGCCAGAAATGAATGAACTGGGAGGAAATGC 780  
QY 781 ACCAGGCTGTTCACATAGATCTAAATCAAACTCTCTCAGACCAATTCGCCACAATTTT 840  
DB 781 ACCAGGCTGTTCACATAGATCTAAATCAAACTCTCTCAGACCAATTCGCCACAATTTT 840  
QY 841 CTTAAAGTAGGGAAACCCCTTATGATAGGTGCAAGCTGTTCATGTGAAACCATGGATT 900  
DB 841 CTTAAAGTAGGGAAACCCCTTATGATAGGTGCAAGCTGTTCATGTGAAACCATGGATT 900  
QY 901 GGGCTACCTGGGAAATTTAGAAAAAACAAGCACTCAGAGGGCAACTACTTTGAGATGAGT 960  
DB 901 GGGCTACCTGGGAAATTTAGAAAAAACAAGCACTCAGAGGGCAACTACTTTGAGATGAGT 960  
QY 961 ACCTATTCAACAAACAGAACTATCATACGGAATCTGTTTGTCTTGTATCATCATGAGTGA 1020  
DB 961 ACCTATTCAACAAACAGAACTATCATACGGAATCTGTTTGTCTTGTATCATCATGAGTGA 1020  
QY 1021 AGAAAACGACACCGGATCTACACTTGTTCCTCTTCAAGCACTCCAGTCAATCAGCTTTG 1080  
DB 1021 AGAAAACGACACCGGATCTACACTTGTTCCTCTTCAAGCACTCCAGTCAATCAGCTTTG 1080  
QY 1081 GTTACCATCTAGGAAAGGGAATTTATAATGCTACCAATTCAGTGAAGATTTATGAATTT 1140  
DB 1081 GTTACCATCTAGGAAAGGGAATTTATAATGCTACCAATTCAGTGAAGATTTATGAATTT 1140  
QY 1141 GACCAATATGAAGAGTTTGTCTTCTGTCAGGTTTAAAGCTTACCCACAATCAGATGT 1200  
DB 1141 GACCAATATGAAGAGTTTGTCTTCTGTCAGGTTTAAAGCTTACCCACAATCAGATGT 1200  
QY 1201 ACCTGGACCTTCTCTCGAAATCAATTTCTTGTGAGCAAAAGGCTTTGATAACGGATAC 1260  
DB 1201 ACCTGGACCTTCTCTCGAAATCAATTTCTTGTGAGCAAAAGGCTTTGATAACGGATAC 1260  
QY 1261 AGCATATCCAAAGTTTGCATATCAAGACACGAGGAGGAGATATATATTTCCATGCGAAA 1320  
DB 1261 AGCATATCCAAAGTTTGCATATCAAGACACGAGGAGGAGATATATATTTCCATGCGAAA 1320  
QY 1321 AATGATGATCCCAATTTTACCAATTTTCAAGCTGATATATGAAGGAAACCTCAAGTG 1380  
DB 1321 AATGATGATCCCAATTTTACCAATTTTCAAGCTGATATATGAAGGAAACCTCAAGTG 1380  
QY 1381 CTGCGAAGACATCGGCAAGTCTAGGCGTCTGTTCTTCGGATGATACCCATTAACCATCT 1440  
DB 1381 CTGCGAAGACATCGGCAAGTCTAGGCGTCTGTTCTTCGGATGATACCCATTAACCATCT 1440

QY 1441 TGGACCTGGAAGAAGTGTTCAGACAAAGTCTCCAACTGCACAGAAAGAGATCAAGAGGA 1500  
DB 1441 TGGACCTGGAAGAAGTGTTCAGACAAAGTCTCCAACTGCACAGAAAGAGATCAAGAGGA 1500  
QY 1501 GTCTGGAATAGAAAGGCTTAAACAGAAAGTGTGTGGACAGTGGGTGTCGAGCAGTACTCTA 1560  
DB 1501 GTCTGGAATAGAAAGGCTTAAACAGAAAGTGTGTGGACAGTGGGTGTCGAGCAGTACTCTA 1560  
QY 1561 AACATGAGTGAAGCCATAAAGGGTCTCTGGTCAAGTGTGTGCATACAAATTCCTTGGC 1620  
DB 1561 AACATGAGTGAAGCCATAAAGGGTCTCTGGTCAAGTGTGTGCATACAAATTCCTTGGC 1620  
QY 1621 ACATCTGTGAGAGCATCTTTTAACTCTCCAGGCCCTTCCCTTTCATCCAAAGCAAC 1680  
DB 1621 ACATCTGTGAGAGCATCTTTTAACTCTCCAGGCCCTTCCCTTTCATCCAAAGCAAC 1680  
QY 1681 ATCTCATTTCTATGCAACAATTTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740  
DB 1681 ATCTCATTTCTATGCAACAATTTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740  
QY 1741 ATTTGTCAAAAGTACAAAGCAATTTTAGGTATGAAAGCCAGCTACAGATGTAAGGTG 1800  
DB 1741 ATTTGTCAAAAGTACAAAGCAATTTTAGGTATGAAAGCCAGCTACAGATGTAAGGTG 1800  
QY 1801 ACCGCTCTCTCAGATAATGAGTACTCTACCTGATGATTTTCAGAGATATGAATATGATCTC 1860  
DB 1801 ACCGCTCTCTCAGATAATGAGTACTCTACCTGATGATTTTCAGAGATATGAATATGATCTC 1860  
QY 1861 AAATGGGAGTTTCCAAGAGAAATTTAGATTTTGGAAAGGTACTAGGATCAGGTGCTTTT 1920  
DB 1861 AAATGGGAGTTTCCAAGAGAAATTTAGATTTTGGAAAGGTACTAGGATCAGGTGCTTTT 1920  
QY 1921 GGAAGAGTGAATGAAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
DB 1921 GGAAGAGTGAATGAAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
QY 1981 GCCGTCAAAATGCTGAAAGAAAGACAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAA 2040  
DB 1981 GCCGTCAAAATGCTGAAAGAAAGACAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAA 2040  
QY 2041 CTCAAGATGATGACCCAGCTGGGAAGCCACAGAAATTTGTGAACTCTGCGGGGCGTGC 2100  
DB 2041 CTCAAGATGATGACCCAGCTGGGAAGCCACAGAAATTTGTGAACTCTGCGGGGCGTGC 2100  
QY 2101 ACCTGTGAGGACCAATTTTACTGATTTTGAATCTGCTATGCTGATCTCTCTCAAC 2160  
DB 2101 ACCTGTGAGGACCAATTTTACTGATTTTGAATCTGCTATGCTGATCTCTCTCAAC 2160  
QY 2161 TATCTAAAGATGAAAGAGAAATTTTCAAGGATTTGGAAGATTTTCAAGGAAACAC 2220  
DB 2161 TATCTAAAGATGAAAGAGAAATTTTCAAGGATTTGGAAGATTTTCAAGGAAACAC 2220  
QY 2221 AATTTGAGTTTACCCCACTTTTCCAAATCTCAGGCTTCTGGAATTTCAATTTTCAAC 2280  
DB 2221 AATTTGAGTTTACCCCACTTTTCCAAATCTCAGGCTTCTGGAATTTCAATTTTCAAC 2280  
QY 2281 GAAGTTCAGATACACCCCGGACTCGGATCAATCTCAGGCTTCTGGAATTTCAATTTTCAAC 2340  
DB 2281 GAAGTTCAGATACACCCCGGACTCGGATCAATCTCAGGCTTCTGGAATTTCAATTTTCAAC 2340  
QY 2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAGGCTGGAAGAGAGGAGGAGCTTGAAT 2400  
DB 2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAGGCTGGAAGAGAGGAGGAGCTTGAAT 2400  
QY 2401 GTGCTTACATTTGAGATCTTCTGCTTTGCTATCAAGTGTGCAAGGAAATGGAATTT 2460  
DB 2401 GTGCTTACATTTGAGATCTTCTGCTTTGCTATCAAGTGTGCAAGGAAATGGAATTT 2460  
QY 2461 CTGGAATTTTAAAGTGTGCTGTTTACAGAGCTGCGCCAGGAAACGCTGCTGTCACCCAC 2520  
DB 2461 CTGGAATTTTAAAGTGTGCTGTTTACAGAGCTGCGCCAGGAAACGCTGCTGTCACCCAC 2520  
QY 2521 GGGAAAGTGTGGAAGATATGATGCTTTGGATTTGCTCGAGATATCATGATGATTTCCAAC 2580

Db	2521	GGAAAGTGTGAAGATATGTGACTTTTGGATTTGGCTCGAGATATCATGAGTGAATCCAAC	2580
Qy	2581	TATGTTGTCAAGGGCAATGCCCGTCTGCTCTGTAATAATGGATGGCCCCCGAAAGCCCTGTTT	2640
Db	2581	TATGTTGTCAAGGGCAATGCCCGTCTGCTCTGTAATAATGGATGGCCCCCGAAAGCCCTGTTT	2640
Qy	2641	GAAGGCATCTACACCAATTAAGAGTGATGCTGTGCATATGGAAATATACTGTGGGAAATC	2700
Db	2641	GAAGGCATCTACACCAATTAAGAGTGATGCTGTGCATATGGAAATATACTGTGGGAAATC	2700
Qy	2701	TTCTCACTTGGTGTGAATCCTTACCCTGGCAATCCGGTTGATGCTAACTTCTACAAACTG	2760
Db	2701	TTCTCACTTGGTGTGAATCCTTACCCTGGCAATCCGGTTGATGCTAACTTCTACAAACTG	2760
Qy	2761	ATTCAAAAATGGATTTAAAAATGATCAGCCCAATTTATGCTACAGAAGAAATATACATTATA	2820
Db	2761	ATTCAAAAATGGATTTAAAAATGATCAGCCCAATTTATGCTACAGAAGAAATATACATTATA	2820
Qy	2821	ATCGAATCCTGCTGGGCTTTTGACTCAAGGAAGACGCCATCTTCCCTAATTTGACTTCG	2880
Db	2821	ATCGAATCCTGCTGGGCTTTTGACTCAAGGAAGACGCCATCTTCCCTAATTTGACTTCG	2880
Qy	2881	TTTTTAGGATGTGAGCTGGCAGATGCAGAAGACGATGTATCAGAAATGTGATGCCCGT	2940
Db	2881	TTTTTAGGATGTGAGCTGGCAGATGCAGAAGACGATGTATCAGAAATGTGATGCCCGT	2940
Qy	2941	GTTCGGAATGCTCTCACACTTACCAGAAACAGGGACCTTTTCAGCAGAGATGGATTTG	3000
Db	2941	GTTCGGAATGCTCTCACACTTACCAGAAACAGGGACCTTTTCAGCAGAGATGGATTTG	3000
Qy	3001	GGGCTACTCTCTCGCAGGCTCAGGTCGAAAGATTCTGAGAGAACCAATTTAGTTTTAAAGG	3060
Db	3001	GGGCTACTCTCTCGCAGGCTCAGGTCGAAAGATTCTGAGAGAACCAATTTAGTTTTAAAGG	3060
Qy	3061	ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAAACAAGATTATTTTCAT	3120
Db	3061	ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAAACAAGATTATTTTCAT	3120
Qy	3121	CAGTAAAAAGAAATCTATTAATCAACTGCTCCTCACCCAGACTTTTCTCTAGAGAGCCGTCT	3180
Db	3121	CAGTAAAAAGAAATCTATTAATCAACTGCTCCTCACCCAGACTTTTCTCTAGAGAGCCGTCT	3180
Qy	3181	GCCTTTACTCTTGTTTTCAAAAGGACCTTTTGTAAAAATCAAAATCATCTGTCACAAGGCAG	3240
Db	3181	GCCTTTACTCTTGTTTTCAAAAGGACCTTTTGTAAAAATCAAAATCATCTGTCACAAGGCAG	3240
Qy	3241	GAGGAGCTGATGAATGAATTTATTTGGAGCAATGATCTGCAATCCAGGCCCTTCTCAGGCCG	3300
Db	3241	GAGGAGCTGATGAATGAATTTATTTGGAGCAATGATCTGCAATCCAGGCCCTTCTCAGGCCG	3300
Qy	3301	GCCTGAGTGAATGTGTACCTGAAAGTACAGTATATTTCTGTAAATAATAAAAACAAGGC	3360
Db	3301	GCCTGAGTGAATGTGTACCTGAAAGTACAGTATATTTCTGTAAATAATAAAAACAAGGC	3360
Qy	3361	ATTTTGTCTAAGGAGAGCTAATATGATTTTTTAAAGTCTATGTTTTTAAAAATAATGTAAA	3420
Db	3361	ATTTTGTCTAAGGAGAGCTAATATGATTTTTTAAAGTCTATGTTTTTAAAAATAATGTAAA	3420
Qy	3421	TTTTTTCAGCTATTTAGTGATATATTTTATGGTGGGAATAAAAAATTTCTACTACAGAAAAA	3480
Db	3421	TTTTTTCAGCTATTTAGTGATATATTTTATGGTGGGAATAAAAAATTTCTACTACAGAAAAA	3480
Qy	3481	AAAAAAAAAAAAAAAAAAAAA 3501	
Db	3481	AAAAAAAAAAAAAAAAAAAAA 3501	

## RESULT 7

RESOLUT /  
PCT-US93-09893-3

PCI-US92-09893-3  
: Sequence 3. Application PC/TUS9209893; sequence 3, application  
: GENERAL INFORMATION:

APPLICANT: Lemischka, Thor R.

Db 301 GAAGTGAATGATCTGCTTCCATCACACTGCAAGTGTGGTCGATGCCCCAGGGAACATT 360  
Qy TCCTGTCTCTGGGTCTTTAAGCAGAGTCCCTGAAATGCCAGGCACACATTTTGAATTACAA 420  
Db TCCTGTCTCTGGGTCTTTAAGCAGAGTCCCTGAAATGCCAGGCACACATTTTGAATTACAA 420  
Qy 421 AACAGAGAGTGTGTTCCATGGTCAATTTGAAATGACAGAAACCCAGCTGGAGATAC 480  
Db AACAGAGAGTGTGTTCCATGGTCAATTTGAAATGACAGAAACCCAGCTGGAGATAC 480  
Qy 481 CTACTTTTATTCAGAGTGAAGTACCAATTTACAAATATTTGTTACAGTGAATAGA 540  
Db CTACTTTTATTCAGAGTGAAGTACCAATTTACAAATATTTGTTACAGTGAATAGA 540  
Qy 541 AATACCTGCTGTATACATTAAGAGACCTTACITTTAGAAAATGGAATAACCCAGAGGCC 600  
Db AATACCTGCTGTATACATTAAGAGACCTTACITTTAGAAAATGGAATAACCCAGAGGCC 600  
Qy 601 CTGGTCTGCATATCTGAGAGCTTCCAGAGCCGATCGTGAATGGGTGCTTTGCGATTCA 660  
Db CTGGTCTGCATATCTGAGAGCTTCCAGAGCCGATCGTGAATGGGTGCTTTGCGATTCA 660  
Qy 661 CAGGGGAAAGCTGTAAGAGAAAGTCCAGCTGTTTAAAAAGGAGGAAAAAGTGCTT 720  
Db CAGGGGAAAGCTGTAAGAGAAAGTCCAGCTGTTTAAAAAGGAGGAAAAAGTGCTT 720  
Qy 721 CATGAATATTTGGGACGACATGAAGTGTGTCGACGAATGAATGAGTGGGAGGGAATGC 780  
Db CATGAATATTTGGGACGACATGAAGTGTGTCGACGAATGAATGAGTGGGAGGGAATGC 780  
Qy 781 ACCAGGCTGTTCACATAGATCTAAATCAAACTCTCAGACCACATTTGCCACAAATATTT 840  
Db ACCAGGCTGTTCACATAGATCTAAATCAAACTCTCAGACCACATTTGCCACAAATATTT 840  
Qy 841 CTTAAAGTAGGGAAACCTTTATGATGAAGTGCAAGCTGTTCATGTGAACCAATGGAATC 900  
Db CTTAAAGTAGGGAAACCTTTATGATGAAGTGCAAGCTGTTCATGTGAACCAATGGAATC 900  
Qy 901 GGCTCACCTGGGAATTAGAAAACAAAGCTCGAGGAGGCACTACTTTGAGATGAGT 960  
Db GGCTCACCTGGGAATTAGAAAACAAAGCTCGAGGAGGCACTACTTTGAGATGAGT 960  
Qy 961 ACCTATTCAAAACAGAACTATGATACGATCTCTGTTGCTTTGATCATCAGTGGCA 1020  
Db ACCTATTCAAAACAGAACTATGATACGATCTCTGTTGCTTTGATCATCAGTGGCA 1020  
Qy 1021 AGAAACGACACCGGATACACTGTTCTCTTCAAAGCATCCAGTCAATCAGCTTTG 1080  
Db AGAAACGACACCGGATACACTGTTCTCTTCAAAGCATCCAGTCAATCAGCTTTG 1080  
Qy 1081 GTTACCATGTAGAAAGGATTTAATATGCTACCAATTTCAAGTGAAGATTTAGAAAT 1140  
Db GTTACCATGTAGAAAGGATTTAATATGCTACCAATTTCAAGTGAAGATTTAGAAAT 1140  
Qy 1141 GACCAATATGAAGATTTGTTTTCTGTCAGGTTTTAAAGCCTTACCCACAAATCAGATG 1200  
Db GACCAATATGAAGATTTGTTTTCTGTCAGGTTTTAAAGCCTTACCCACAAATCAGATG 1200  
Qy 1201 AGTGGACCTTCTCTCGAAAATCAITTTCTTGTGAGCAAAAGGCTTTGATAACGGATAC 1260  
Db AGTGGACCTTCTCTCGAAAATCAITTTCTTGTGAGCAAAAGGCTTTGATAACGGATAC 1260  
Qy 1261 AGCATATCAAGTTTTGCAATATAGACACCGGAGGAATATATATTTCCATGAGAA 1320  
Db AGCATATCAAGTTTTGCAATATAGACACCGGAGGAATATATATTTCCATGAGAA 1320  
Qy 1321 AATGATGAGCCCAATTTACCAAAATGTTACCGCTGAAATAGAGGAAACCTCAAGTG 1380  
Db AATGATGAGCCCAATTTTACCAAAATGTTACCGCTGAAATAGAGGAAACCTCAAGTG 1380  
Qy 1381 CTGCGAAGAGCATCGGCAAGTCAAGGCTGCTGTTTCTCGATGATACCCATACCATCT 1440  
Db CTGCGAAGAGCATCGGCAAGTCAAGGCTGCTGTTTCTCGATGATACCCATACCATCT 1440

Qy 1441 TGGACCTGGAGAGTGTTCAGACAAAGTCTCCCAACTGCCACAGAGAGATCACAGAAAGA 1500  
Db TGGACCTGGAGAGTGTTCAGACAAAGTCTCCCAACTGCCACAGAGAGATCACAGAAAGA 1500  
Qy 1501 GTCTGGAATAGAAAGGCTTAACAGAAAAAGTGTTCGACAGTGGGTGTCGAGCAGTACTCTA 1560  
Db GTCTGGAATAGAAAGGCTTAACAGAAAAAGTGTTCGACAGTGGGTGTCGAGCAGTACTCTA 1560  
Qy 1561 AACATGATGAAGCCATTAAGAGGTTCTTGGTCAAGTGTGTCATACAATTCCTTGGC 1620  
Db AACATGATGAAGCCATTAAGAGGTTCTTGGTCAAGTGTGTCATACAATTCCTTGGC 1620  
Qy 1621 ACATCTTGTGAGAGCTCTTTTAACTCTCCAGGCCCTTCCCTTTTCAATCCAGAGAAC 1680  
Db ACATCTTGTGAGAGCTCTTTTAACTCTCCAGGCCCTTCCCTTTTCAATCCAGAGAAC 1680  
Qy 1681 ATCTCATCTTATGCAACAATTTGGTGTGTTCTCTCTCTTCAATGTCGTTTAACTCTGTA 1740  
Db ATCTCATCTTATGCAACAATTTGGTGTGTTCTCTCTCTTCAATGTCGTTTAACTCTGTA 1740  
Qy 1741 ATTTGTCAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
Db ATTTGTCAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
Qy 1801 ACCGGTCTCTCAGATATGATGTTCTTCTAGTGTGTTTCAAGAAATATGAATATGATCTC 1860  
Db ACCGGTCTCTCAGATATGATGTTCTTCTAGTGTGTTTCAAGAAATATGAATATGATCTC 1860  
Qy 1861 AAATGGGAGTTTCCAGAGAAATTTAGAGTTTGGGAAAGTACTAGGATCAGGTGCTTTT 1920  
Db AAATGGGAGTTTCCAGAGAAATTTAGAGTTTGGGAAAGTACTAGGATCAGGTGCTTTT 1920  
Qy 1921 GGAAGATGATGAACGCAACAGCTTTATGAAATAGCAAAAACAGAGTCTCAATCCAGTT 1980  
Db GGAAGATGATGAACGCAACAGCTTTATGAAATAGCAAAAACAGAGTCTCAATCCAGTT 1980  
Qy 1981 GCGTCAAAATGCTGAAAGAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCAAG 2040  
Db GCGTCAAAATGCTGAAAGAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCAAG 2040  
Qy 2041 CTCAAGATGATGACCCAGCTGGGAAGCCAGAAATTTTGAACCTTGGGAGGCTGTC 2100  
Db CTCAAGATGATGACCCAGCTGGGAAGCCAGAAATTTTGAACCTTGGGAGGCTGTC 2100  
Qy 2101 ACACTGTCAAGCAAAATTTACTTTGAAATCTGTTGTAATCTGTTGATCTTCTCAAC 2160  
Db ACACTGTCAAGCAAAATTTACTTTGAAATCTGTTGTAATCTGTTGATCTTCTCAAC 2160  
Qy 2161 TATCTAAGAGTAAAGAGAAAAATTTCAAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
Db TATCTAAGAGTAAAGAGAAAAATTTCAAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
Qy 2221 AATTTCAAGTTTTTACCCCACTTCCATCAATCCAAATTCAGGCTGCTGGTTCAGG 2280  
Db AATTTCAAGTTTTTACCCCACTTCCATCAATTCAGGCTGCTGGTTCAGG 2280  
Qy 2281 GAAATTCAGATACACCCGAGCTCGGATCAAACTCTCAGGCTTCAATTCAGGATTCAT 2340  
Db GAAATTCAGATACACCCGAGCTCGGATCAAACTCTCAGGCTTCAATTCAGGATTCAT 2340  
Qy 2341 TCTGAAGATGAAATTTGAATATGAAAAACCAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400  
Db TCTGAAGATGAAATTTGAATATGAAAAACCAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400  
Qy 2401 GTGCTTACATTTGAGATCTTCTTTGCAATATCAAGTTGCAAGTGGAAATTTT 2460  
Db GTGCTTACATTTGAGATCTTCTTTGCAATATCAAGTTGCAAGTGGAAATTTT 2460  
Qy 2461 CTGGAATTTAAGTCGTGTTTCAAGAGACCTGGCCGCCAGGAACGCTGTTGTCAACCCAC 2520  
Db CTGGAATTTAAGTCGTGTTTCAAGAGACCTGGCCGCCAGGAACGCTGTTGTCAACCCAC 2520

Qy	2521	GGGAAAGTGGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTGATCCCAAC	2580
Db	2521	GGGAAAGTGGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTGATCCCAAC	2580
Qy	2581	TATGTTGCTCAGGGCAATGCGCGTCTGCCTGTAAAAATGGATGGCCGCCGGAAGACCTGTTTT	2640
Db	2581	TATGTTGCTCAGGGCAATGCGCGTCTGCCTGTAAAAATGGATGGCCGCCGGAAGACCTGTTTT	2640
Qy	2641	GAAGGCATCTACACCATTAAGAGTGATGTCTGTGTCATATGGAATATATCTGTGTGGGAAATC	2700
Db	2641	GAAGGCATCTACACCATTAAGAGTGATGTCTGTGTCATATGGAATATATCTGTGTGGGAAATC	2700
Qy	2701	TTCTCACTTGGTGTGAATCCTTACCCTGGCATTCGGTTCAGCTAACTTCTTCAAACTG	2760
Db	2701	TTCTCACTTGGTGTGAATCCTTACCCTGGCATTCGGTTCAGCTAACTTCTTCAAACTG	2760
Qy	2761	ATTCMAAATGGATTTAAAAATGGATCAGCCATTTTATGCTACAGAAGAAATATACATTATA	2820
Db	2761	ATTCMAAATGGATTTAAAAATGGATCAGCCATTTTATGCTACAGAAGAAATATACATTATA	2820
Qy	2821	ATGCAATCTCTGTGGGCTTTTGACTCAAGGAACGGCCATCCTTCCCTAAATTTGACTTCG	2880
Db	2821	ATGCAATCTCTGTGGGCTTTTGACTCAAGGAACGGCCATCCTTCCCTAAATTTGACTTCG	2880
Qy	2881	TTTTTTAGGATGTCTCAGCTGGCAGATGTCAGAAAGACGGATGTATCAGAATGTGGATGGCCGT	2940
Db	2881	TTTTTTAGGATGTCTCAGCTGGCAGATGTCAGAAAGACGGATGTATCAGAATGTGGATGGCCGT	2940
Qy	2941	GTTTTGGAAATGTCTCACACCTACAAAAACAGGCGACCTTTTCAGCAGAGATGGATTTG	3000
Db	2941	GTTTTGGAAATGTCTCACACCTACAAAAACAGGCGACCTTTTCAGCAGAGATGGATTTG	3000
Qy	3001	GGGCTACTCTCTCGGAGGCTCAGGTGGAAGATTGCTAGAGGAACAATTTAGTTTTAAGG	3060
Db	3001	GGGCTACTCTCTCGGAGGCTCAGGTGGAAGATTGCTAGAGGAACAATTTAGTTTTAAGG	3060
Qy	3061	ACTTTCATCCCTCCACCTATCCTTAAACAGGCTGTAGATTTACCAAAACAAGATTAATTTTCAT	3120
Db	3061	ACTTTCATCCCTCCACCTATCCTTAAACAGGCTGTAGATTTACCAAAACAAGATTAATTTTCAT	3120
Qy	3121	CAC'TAAAGAAAAATCTATATTCAACTGCTGCTTCCACAGACTTTTCTCTAGAGCCGTCT	3180
Db	3121	CAC'TAAAGAAAAATCTATATTCAACTGCTGCTTCCACAGACTTTTCTCTAGAGCCGTCT	3180
Qy	3181	CGGTTTACTCTTGTTTTCAAAGGACATTTGTAAATCAAATCATCTGTCACAAAGGCAG	3240
Db	3181	CGGTTTACTCTTGTTTTCAAAGGACATTTGTAAATCAAATCATCTGTCACAAAGGCAG	3240
Qy	3241	GAGGAGCTGATAATGAACCTTTATTTGGAGCATTTGATCTGCATCCAAAGGCCCTTCTCAGGCCG	3300
Db	3241	GAGGAGCTGATAATGAACCTTTATTTGGAGCATTTGATCTGCATCCAAAGGCCCTTCTCAGGCCG	3300
Qy	3301	GC'TTGAAGTAAATGTGTACCTGGAAGTACATATATCTTGTAAATACATATAAAACAAAGC	3360
Db	3301	GC'TTGAAGTAAATGTGTACCTGGAAGTACATATATCTTGTAAATACATATAAAACAAAGC	3360
Qy	3361	ATTTTTGCTTAGGAGAGCTAATATGATTTTTTTTAACTCTATGTTTTTAAAAATAATGTAAA	3420
Db	3361	ATTTTTGCTTAGGAGAGCTAATATGATTTTTTTTAACTCTATGTTTTTAAAAATAATGTAAA	3420
Qy	3421	TTTTTTCAGCTATTTTAGTGATATATTTTATGGGTGGGAATAAAATTTCTTACTACAGAAAA	3480
Db	3421	TTTTTTCAGCTATTTTAGTGATATATTTTATGGGTGGGAATAAAATTTCTTACTACAGAAAA	3480
Qy	3481	AAAAAAAAAAAAAAAAAAAAA 3501	
Db	3481	AAAAAAAAAAAAAAAAAAAAA 3501	

GENERAL INFORMATION:  
APPLICANT: Bennett, Brian D.  
APPLICANT: Broz, Susan D.  
APPLICANT: Matthews, William C.  
APPLICANT: Zeigler, Francis C.  
TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES THE  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/222,299  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Hasak, Janet E.  
REGISTRATION NUMBER: 28,616  
REFERENCE/DOCKET NUMBER: 879  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1896  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3475 bases  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

Query Match	99.3%; Score 3475; DB 1; Length 3475;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 3475; Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
QY 1	CGAGGCGGCATCGAGGGCTGGGCGCGCGCCCTGGGGGACCCCGGGCTCCGAGGCCATG 60
Db 1	CGAGGCGGCATCGAGGGCTGGGCGCGCGCCCTGGGGGACCCCGGGCTCCGAGGCCATG 60
QY 61	CCGCGCTTGGCGCGCGACGCGGGCACCGTGCCGCTGCTCGTTGTTTTTTCTGCAATGATA 120
Db 61	CCGCGCTTGGCGCGCGACGCGGGCACCGTGCCGCTGCTCGTTGTTTTTTCTGCAATGATA 120
QY 121	TTTGGGACTATTACAATCAAGATCTGCCTGATCAAGTGCTGTTTAAATCAATCATAAG 180
Db 121	TTTGGGACTATTACAATCAAGATCTGCCTGATCAAGTGCTGTTTAAATCAATCATAAG 180
QY 181	AACAAATGATTTCATCAGTGGGGAAAGTCATCATCATATCCATGGTGATCAGAATCCCCGGAA 240
Db 181	AACAAATGATTTCATCAGTGGGGAAAGTCATCATCATATCCATGGTGATCAGAATCCCCGGAA 240
QY 241	GACCTCGGGTGCTGGTTGAGACCCAGAGCTCAGGGACAGTGTAAGAGCTGGCGCTGTG 300
Db 241	GACCTCGGGTGCTGGTTGAGACCCAGAGCTCAGGGACAGTGTAAGAGCTGGCGCTGTG 300
QY 301	GAAGTGATGTATCTGCTTCCATCACACTGCCAAGTGCTGGTCGATGCCCGAGGGAAACATT 360
Db 301	GAAGTGATGTATCTGCTTCCATCACACTGGAAAGTGCTGGTCGATGCCCGAGGGAAACATT 360
QY 361	TCCTGTCTCTGGGCTTTTAAAGCAAGCTCCCTGAAATTTGCCAGGCCACATTTTGTATTACAA 420
Db 361	TCCTGTCTCTGGGCTTTTAAAGCAAGCTCCCTGAAATTTGCCAGGCCACATTTTGTATTACAA 420

421 AACAGAGGAGTGTGTTCCATGTCATTTTGAATAACAGAGAAACCCAGCTGAGAAATAC 480  
Db  
421 AACAGAGGAGTGTGTTCCATGTCATTTTGAATAACAGAGAAACCCAGCTGAGAAATAC 480  
Qy  
481 CTACTTTTATTACAGAGTGAAGCTACCAATTAACAAATATGTTTACAGTGAGTATAAGA 540  
Db  
481 CTACTTTTATTACAGAGTGAAGCTACCAATTAACAAATATGTTTACAGTGAGTATAAGA 540  
Qy  
541 AATACCTCTCTTACACATTAAGAGACCTTACTTTAGAAAATGAGAAACCCAGGACGCC 600  
Db  
541 AATACCTCTCTTACACATTAAGAGACCTTACTTTAGAAAATGAGAAACCCAGGACGCC 600  
Qy  
601 CTGCTCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGAATGSGTGTCTTTGCGATTCA 660  
Db  
601 CTGCTCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGAATGSGTGTCTTTGCGATTCA 660  
Qy  
661 CAGGGGGAAGCTGTAAAGAGAAAGTCCAGCTGTTGTTTAAAGAGGAAAAGTGCCT 720  
Db  
661 CAGGGGGAAGCTGTAAAGAGAAAGTCCAGCTGTTGTTTAAAGAGGAAAAGTGCCT 720  
Qy  
721 CATGAATATTATTTGGGACGACATTAAGTGTCTGTCAGCAAAATGAATGAACTGGGCGAGGAATGC 780  
Db  
721 CATGAATATTATTTGGGACGACATTAAGTGTCTGTCAGCAAAATGAATGAACTGGGCGAGGAATGC 780  
Qy  
781 ACCAGGCTGTTCACAATAGATCTTAATCAAACTCTCTCAGACACATTTGCCACAATTAATTT 840  
Db  
781 ACCAGGCTGTTCACAATAGATCTTAATCAAACTCTCTCAGACACATTTGCCACAATTAATTT 840  
Qy  
841 CTTAAAGTAGGGGAACCCCTTATGATAGAGTGAAGCTGTTTCAATGTAACCATGATGATTC 900  
Db  
841 CTTAAAGTAGGGGAACCCCTTATGATAGAGTGAAGCTGTTTCAATGTAACCATGATGATTC 900  
Qy  
901 GGGCTCACCTGGGAATTAGAAAACAAAGCACTCGAGGAGGGCAACTATTGAGATGAGT 960  
Db  
901 GGGCTCACCTGGGAATTAGAAAACAAAGCACTCGAGGAGGGCAACTATTGAGATGAGT 960  
Qy  
961 ACCTATTCAACAAAGACTATGATACGGATCTGTTTCTGTTTCTGTTTATCATCAGTGGCA 1020  
Db  
961 ACCTATTCAACAAAGACTATGATACGGATCTGTTTCTGTTTCTGTTTATCATCAGTGGCA 1020  
Qy  
1021 AGAAACGACACCGGATCTACACTTGTCTCTTCAAGACATCCCAAGTCAATCAGCTTTG 1080  
Db  
1021 AGAAACGACACCGGATCTACACTTGTCTCTTCAAGACATCCCAAGTCAATCAGCTTTG 1080  
Qy  
1081 GTTACCATCTGAGGAAAGGGATTTATAATGCTACCAATTTCAAGTGAAGATTAAGAAAT 1140  
Db  
1081 GTTACCATCTGAGGAAAGGGATTTATAATGCTACCAATTTCAAGTGAAGATTAAGAAAT 1140  
Qy  
1141 GACCAATATGAAGATTTTGTGTTTCTGTCAGGTTTAAAGCCTACCCACAAATCAGATGT 1200  
Db  
1141 GACCAATATGAAGATTTTGTGTTTCTGTCAGGTTTAAAGCCTACCCACAAATCAGATGT 1200  
Qy  
1201 ACGTGGACCTCTCTCGAAAATCAATTCCTCTGTGAGCAAAAGGGTCTTGATACCGGATAC 1260  
Db  
1201 ACGTGGACCTCTCTCGAAAATCAATTCCTCTGTGAGCAAAAGGGTCTTGATACCGGATAC 1260  
Qy  
1261 AGCATATCCAAATTTTGAATCATTAAGCACACAGCCAGGAGATATATATTTCCATGAGAA 1320  
Db  
1261 AGCATATCCAAATTTTGAATCATTAAGCACACAGCCAGGAGATATATATTTCCATGAGAA 1320  
Qy  
1321 AATGATATGCCAAATTTTCAAAAATGTTTACGCTGAATATTAAGAGGAAACCTCAAGTG 1380  
Db  
1321 AATGATATGCCAAATTTTCAAAAATGTTTACGCTGAATATTAAGAGGAAACCTCAAGTG 1380  
Qy  
1381 CTCGAGAGGATCGGCAAGTCAAGGCTCTGTTTCTCGGATGGATACCCATTACATCT 1440  
Db  
1381 CTCGAGAGGATCGGCAAGTCAAGGCTCTGTTTCTCGGATGGATACCCATTACATCT 1440  
Qy  
1441 TGGACCTGGAAGAGTGTTCAGACAAAGTCTCCAACTGCAAGAGAGATCAAGAGGA 1500  
Db  
1441 TGGACCTGGAAGAGTGTTCAGACAAAGTCTCCAACTGCAAGAGAGATCAAGAGGA 1500

1501 GTCTGGAATAGAAAGCGCTAAACAGAAAAGTGTGGAACAGTGGGTGTCGAGCAGTACTCTA 1560  
Db  
1501 GTCTGGAATAGAAAGCGCTAAACAGAAAAGTGTGGAACAGTGGGTGTCGAGCAGTACTCTA 1560  
Qy  
1561 AACATCAGTGAAAGCCATAAAGGGTTCCTGGTCAAGTGTGTGCATACAATTCCTTGGC 1620  
Db  
1561 AACATCAGTGAAAGCCATAAAGGGTTCCTGGTCAAGTGTGTGCATACAATTCCTTGGC 1620  
Qy  
1621 ACATCTTGTGAGACGATCCCTTTTAACTCTCAGGCCCCCTTCCCTTTTCAATCAAGACAAC 1680  
Db  
1621 ACATCTTGTGAGACGATCCCTTTTAACTCTCAGGCCCCCTTCCCTTTTCAATCAAGACAAC 1680  
Qy  
1681 ATCTCATCTATGCAACAATTCGTGTGTTCTCTCTCTTCAATTCGTGCTTTTAAACCTGCTA 1740  
Db  
1681 ATCTCATCTATGCAACAATTCGTGTGTTCTCTCTCTTCAATTCGTGCTTTTAAACCTGCTA 1740  
Qy  
1741 ATTTGTCACAAGTACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGGTG 1800  
Db  
1741 ATTTGTCACAAGTACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGGTG 1800  
Qy  
1801 ACCGGCTCTCAGATAATGAGTACTTCTACGTTGATTTTCAAGAGATATGAAATGATATCTC 1860  
Db  
1801 ACCGGCTCTCAGATAATGAGTACTTCTACGTTGATTTTCAAGAGATATGAAATGATATCTC 1860  
Qy  
1861 AAATCGGAGTTCCTCAAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920  
Db  
1861 AAATCGGAGTTCCTCAAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920  
Qy  
1921 GGAAGTGTGTAAGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
Db  
1921 GGAAGTGTGTAAGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
Qy  
1981 GCGGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAA 2040  
Db  
1981 GCGGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAA 2040  
Qy  
2041 CTCAGATGATGACCCAGCTGGGAAGCCAGAGAAATATTTGTAACCTGCTGGGGGGTGC 2100  
Db  
2041 CTCAGATGATGACCCAGCTGGGAAGCCAGAGAAATATTTGTAACCTGCTGGGGGGTGC 2100  
Qy  
2101 ACACGTCTCAGGACCAATTTTACTTGAATTTTGAATCTGCTATGCTGATCTTCTCAAC 2160  
Db  
2101 ACACGTCTCAGGACCAATTTTACTTGAATTTTGAATCTGCTATGCTGATCTTCTCAAC 2160  
Qy  
2161 TATCTAAGAAGTAAAGAGAAAATTTTCAAGGATTTGACAGAGATTTTCAAGGAACAC 2220  
Db  
2161 TATCTAAGAAGTAAAGAGAAAATTTTCAAGGATTTGACAGAGATTTTCAAGGAACAC 2220  
Qy  
2221 AATTTCAAGTTTACCCCACTTTTCAATCAATCCAAATTTCCAGCATGCTGTTCAAGA 2280  
Db  
2221 AATTTCAAGTTTACCCCACTTTTCAATCAATCCAAATTTCCAGCATGCTGTTCAAGA 2280  
Qy  
2281 GAAGTTCAAGTACACCCGAGCTCGGATCAAAATCTCAGGCTTCAAGGAAATTCATTTCAC 2340  
Db  
2281 GAAGTTCAAGTACACCCGAGCTCGGATCAAAATCTCAGGCTTCAAGGAAATTCATTTCAC 2340  
Qy  
2341 TCTGAAGATGAATTAATGAATATGAATAAGGCTGGAAGAGAGAGGAGACTTGAAT 2400  
Db  
2341 TCTGAAGATGAATTAATGAATATGAATAAGGCTGGAAGAGAGAGGAGACTTGAAT 2400  
Qy  
2401 GTGCTTACATTTGAAGATCTTTTGTGTTTGCATATCAAGTTGCCAAGGAATGAAATTT 2460  
Db  
2401 GTGCTTACATTTGAAGATCTTTTGTGTTTGCATATCAAGTTGCCAAGGAATGAAATTT 2460  
Qy  
2461 CTGGAATTTAAGTGTGTGTTTCAAGACCTGGCCCGCAGGAACTGCTTGTCCACCCAC 2520  
Db  
2461 CTGGAATTTAAGTGTGTGTTTCAAGACCTGGCCCGCAGGAACTGCTTGTCCACCCAC 2520  
Qy  
2521 GGGAAAGTGTGAAGATATGATCTTTGATTTGGTCCAGATATCATGAGTATTTCCAC 2580  
Db  
2521 GGGAAAGTGTGAAGATATGATCTTTGATTTGGTCCAGATATCATGAGTATTTCCAC 2580  
Qy  
2581 TATGTTGTGAGGGCAATGCGCTCTGCTGTAAAAATGGATGGCCCCCGAAGACCTGTTT 2640







QY 541 AATACCTGCTTTACATTAAGAGACCTTACTTTAGAAAAATGAAAAACAGAGCC 600  
DB 541 AATACCTGCTTTACATTAAGAGACCTTACTTTAGAAAAATGAAAAACAGAGCC 600  
QY 601 CTGCTCTGCATATCTGAGAGGTTCCAGAGCGATCGTGGAAATGGGTCTTTGCGATTCA 660  
DB 601 CTGCTCTGCATATCTGAGAGGTTCCAGAGCGATCGTGGAAATGGGTCTTTGCGATTCA 660  
QY 661 CAGGGGAAAGCTGTAAGAGAAAGTCCAGCTGTTGTTAAAAAGGAGGAAAAAGTCTT 720  
DB 661 CAGGGGAAAGCTGTAAGAGAAAGTCCAGCTGTTGTTAAAAAGGAGGAAAAAGTCTT 720  
QY 721 CATGAATTTATTTGGGACGACATTAAGGTCTGTGCCAGAAATGAATCGGCGAGGAATGC 780  
DB 721 CATGAATTTATTTGGGACGACATTAAGGTCTGTGCCAGAAATGAATCGGCGAGGAATGC 780  
QY 781 ACCAGGCTGTTTCAATATAGATCTTAAATCAAACTCCTCAGACCAATTCGCCAATTTT 840  
DB 781 ACCAGGCTGTTTCAATATAGATCTTAAATCAAACTCCTCAGACCAATTCGCCAATTTT 840  
QY 841 CTTAAAGTAGGGAAACCTTTATGATTAAGGTGCAAGCTGTTCAATGCAACATGGATTTC 900  
DB 841 CTTAAAGTAGGGAAACCTTTATGATTAAGGTGCAAGCTGTTCAATGCAACATGGATTTC 900  
QY 901 GGGCTCACCTGGGAAATTAGAAAACAAAAGCACTCGAGGAGGCAACTACTTTGAGATGAT 960  
DB 901 GGGCTCACCTGGGAAATTAGAAAACAAAAGCACTCGAGGAGGCAACTACTTTGAGATGAT 960  
QY 961 ACCTATTCAACAAACAGAACTATGATACGGATCTGTTGCTTTGTATATCAATGAGTGC 1020  
DB 961 ACCTATTCAACAAACAGAACTATGATACGGATCTGTTGCTTTGTATATCAATGAGTGC 1020  
QY 1021 AGAAACGACACCGGATACACATGTTGCTCTTCAAAAGCATCCAGTCAATGAGTTC 1080  
DB 1021 AGAAACGACACCGGATACACATGTTGCTCTTCAAAAGCATCCAGTCAATGAGTTC 1080  
QY 1081 GTTACCATCTGAGAAAGGATTTATTAATGCTACCAATTCAGTGAAGATTTGAAATT 1140  
DB 1081 GTTACCATCTGAGAAAGGATTTATTAATGCTACCAATTCAGTGAAGATTTGAAATT 1140  
QY 1141 GACCAATATGAAGAGTTTGTGTTCTGTAGGTTTAAAGCTTACCCCAAAATCAGATGT 1200  
DB 1141 GACCAATATGAAGAGTTTGTGTTCTGTAGGTTTAAAGCTTACCCCAAAATCAGATGT 1200  
QY 1201 ACGTGGACCTTCTCTGAAAAATCAATTCCTGAGCAAAAGGCTTTGATAACGGATAC 1260  
DB 1201 ACGTGGACCTTCTCTGAAAAATCAATTCCTGAGCAAAAGGCTTTGATAACGGATAC 1260  
QY 1261 AGCATATCCAAGTTTGGCAATCATAGCACAGCAGAGAGAAATATATATTCATGCGAAA 1320  
DB 1261 AGCATATCCAAGTTTGGCAATCATAGCACAGCAGAGAGAAATATATATTCATGCGAAA 1320  
QY 1321 AATGATGATGCCAATTTTACCAAAATGTTTCAAGCTGAAATTAAGAGGAAACCTCAAGT 1380  
DB 1321 AATGATGATGCCAATTTTACCAAAATGTTTCAAGCTGAAATTAAGAGGAAACCTCAAGT 1380  
QY 1381 CTCGAGAGCATCGGCAAGTCAAGCGCTCTGTTTCTCGGATGGATACCCATTACCACT 1440  
DB 1381 CTCGAGAGCATCGGCAAGTCAAGCGCTCTGTTTCTCGGATGGATACCCATTACCACT 1440  
QY 1441 TGGACCTGGAAGAGTGTTCAGACAAAGTCTCCCACTGCAAGAGAGATCAGAGAGGA 1500  
DB 1441 TGGACCTGGAAGAGTGTTCAGACAAAGTCTCCCACTGCAAGAGAGATCAGAGAGGA 1500  
QY 1501 GTCTGGAATAGAAAGGCTTAAACAGAAAGTGTTCGACAGTGGGTTCGAGCAGTACTCTA 1560  
DB 1501 GTCTGGAATAGAAAGGCTTAAACAGAAAGTGTTCGACAGTGGGTTCGAGCAGTACTCTA 1560  
QY 1561 AACATGAGTGAAGCATAAAAGGTTCTGTTGCAAGTCTGTGCATACAAATTCCTTTGGC 1620  
DB 1561 AACATGAGTGAAGCATAAAAGGTTCTGTTGCAAGTCTGTGCATACAAATTCCTTTGGC 1620  
QY 1621 ACATCTTTGTGAGACGATCCTTTTAACTCTCCAGGCCCCCTTCCCTTTTCATCCAAAGACAAC 1680

DB 1621 ACATCTTTGTGAGACGATCCTTTTAACTCTCCAGGCCCCCTTCCCTTTTCATCCAAAGACAAC 1680  
QY 1681 ATCTCATTTATGCAACAAATTTGGTGTGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740  
DB 1681 ATCTCATTTATGCAACAAATTTGGTGTGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740  
QY 1741 ATTTGTGCAGAGTCAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGTTGACAGGTG 1800  
DB 1741 ATTTGTGCAGAGTCAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGTTGACAGGTG 1800  
QY 1801 ACCGGCTCTCAGATAATGAGTACTCTACGTTGATTTTTCAGAGAAATATGAATATGATCTC 1860  
DB 1801 ACCGGCTCTCAGATAATGAGTACTCTACGTTGATTTTTCAGAGAAATATGAATATGATCTC 1860  
QY 1861 AAATGGGAGTTTCCAGAGAAAAATTTAGAGTTTGGGAAAGTACTAGGATCAGGTCCTTTT 1920  
DB 1861 AAATGGGAGTTTCCAGAGAAAAATTTAGAGTTTGGGAAAGTACTAGGATCAGGTCCTTTT 1920  
QY 1921 GGAAAGTGTATGAACGCAACAGCTTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
DB 1921 GGAAAGTGTATGAACGCAACAGCTTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
QY 1981 GCCGTCAAAATGCTGAAAGAAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAA 2040  
DB 1981 GCCGTCAAAATGCTGAAAGAAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAA 2040  
QY 2041 CTCAGATGATGACCCAGCTGGGAAAGCCAGAGAAATTTGTGAACCTCTGCTGGGGCGTGC 2100  
DB 2041 CTCAGATGATGACCCAGCTGGGAAAGCCAGAGAAATTTGTGAACCTCTGCTGGGGCGTGC 2100  
QY 2101 ACATGTCAGAGCAAAATTTACTTCTGATTTTTCGAATTAATGTTGCTATGTTGATCTCTCAAC 2160  
DB 2101 ACATGTCAGAGCAAAATTTACTTCTGATTTTTCGAATTAATGTTGCTATGTTGATCTCTCAAC 2160  
QY 2161 TATCTAAGAGTAAAGAGAAAAATTTTCAAGGACTTTCGACAGAGATTTTCAAGGAAACAC 2220  
DB 2161 TATCTAAGAGTAAAGAGAAAAATTTTCAAGGACTTTCGACAGAGATTTTCAAGGAAACAC 2220  
QY 2221 AAATTCAGTTTTTACCCCACTTTTCCAATCACATCCAAATTCAGCATCTCGGTTTCAAGA 2280  
DB 2221 AAATTCAGTTTTTACCCCACTTTTCCAATCACATCCAAATTCAGCATCTCGGTTTCAAGA 2280  
QY 2281 GAAGTTCAGATACACCCGAGCTCGGATCAATCTCAGGGCTTCATGGGAAATTCATTTCAAC 2340  
DB 2281 GAAGTTCAGATACACCCGAGCTCGGATCAATCTCAGGGCTTCATGGGAAATTCATTTCAAC 2340  
QY 2341 TCTGAAGATGAAATTTGAATATGAAGAACCAAAAGGCTGGAAGAGAGAGGAGACTTGAAT 2400  
DB 2341 TCTGAAGATGAAATTTGAATATGAAGAACCAAAAGGCTGGAAGAGAGAGGAGACTTGAAT 2400  
QY 2401 GTGCTTACATTTGAAGATCTTTTGTGTTTGCATATCAAGTTGCAAGGAAATGGAATTT 2460  
DB 2401 GTGCTTACATTTGAAGATCTTTTGTGTTTGCATATCAAGTTGCAAGGAAATGGAATTT 2460  
QY 2461 CTGAAATTTAAGTCTGTGTTTTCAGAGACTGCGCGCCAGGAAACGTCCTTGTCAACCAAC 2520  
DB 2461 CTGAAATTTAAGTCTGTGTTTTCAGAGACTGCGCGCCAGGAAACGTCCTTGTCAACCAAC 2520  
QY 2521 GGGAAAGTGGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTGAATTCACAC 2580  
DB 2521 GGGAAAGTGGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTGAATTCACAC 2580  
QY 2581 TATGTTGTGAGGGCAATGCCCGTCTGCTCTGTAATAATGGAATGGAATGGAATGGAATTCACAC 2640  
DB 2581 TATGTTGTGAGGGCAATGCCCGTCTGCTCTGTAATAATGGAATGGAATGGAATGGAATTCACAC 2640  
QY 2641 GAAGGCAATCACCAATTAAGAGTGAATGTTGCTGTTGATGATGATGATGATGATGATGATGAT 2700  
DB 2641 GAAGGCAATCACCAATTAAGAGTGAATGTTGCTGTTGATGATGATGATGATGATGATGATGAT 2700  
QY 2701 TTCTCATTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2760

2701	TTCTCACTTGGTGTGAATCCTTACCCTGGCATCCCGTGTGATGCTAACTTCTACAACCTG	2760
2761	ATTCAAAATGGATTTAAAAATGGATCAGCCATTTTATGCTACAGAGAAATATACATTATA	2820
2761	ATTCAAAATGGATTTAAAAATGGATCAGCCATTTTATGCTACAGAGAAATATACATTATA	2820
2821	ATGCAATCTCTGCTGGGCTTTTGATCTAAGGAAACGGCCATCCTTCCCTAAATTGACTTCG	2880
2821	ATGCAATCTCTGCTGGGCTTTTGATCTAAGGAAACGGCCATCCTTCCCTAAATTGACTTCG	2880
2881	TTTTTAGGATGTCAGCTGGCAGATGCGAAGAAGCGGATGATCAGAAATGTGGATGCGCGT	2940
2881	TTTTTAGGATGTCAGCTGGCAGATGCGAAGAAGCGGATGATCAGAAATGTGGATGCGCGT	2940
2941	GTTTTCGGAATGTCCTCACACCTACCAAAAACAGCGGACCTTTTCAGCAGAGAGATGGAATTG	3000
2941	GTTTTCGGAATGTCCTCACACCTACCAAAAACAGCGGACCTTTTCAGCAGAGAGATGGAATTG	3000
3001	GGGCTACTCTCTCGGAGGCTCAGGTCGAAGATTCGTAGAGGAACAATTTAGTTTTAAGG	3060
3001	GGGCTACTCTCTCGGAGGCTCAGGTCGAAGATTCGTAGAGGAACAATTTAGTTTTAAGG	3060
3061	ACTTCATCCCTCCACCTATCCCTAAACAGGCTGTAGATTACCAAAAACAGGAATTAATTTTCAT	3120
3061	ACTTCATCCCTCCACCTATCCCTAAACAGGCTGTAGATTACCAAAAACAGGAATTAATTTTCAT	3120
3121	CACATAAAGAAAATCTATTATCAACTGCTGCTCTCACAGAGACTTTTCTCTAGAAGCGTCT	3180
3121	CACATAAAGAAAATCTATTATCAACTGCTGCTCTCACAGAGACTTTTCTCTAGAAGCGTCT	3180
3181	GCCTTTACTCTGTTTTCAAAAGGACTTTTGTGTAATAATCAATCATCTGTCAAGGCGAG	3240
3181	GCCTTTACTCTGTTTTCAAAAGGACTTTTGTGTAATAATCAATCATCTGTCAAGGCGAG	3240
3241	GAGGAGCTGATATGAATCTTATTTGGAGCAATTCATCTGCATCCAAGGCTTCTCAGGCGG	3300
3241	GAGGAGCTGATATGAATCTTATTTGGAGCAATTCATCTGCATCCAAGGCTTCTCAGGCGG	3300
3301	GCCTTAGTGAATTTGTACCTGGAAGTACAGTATATCTTGTAAATACATAAAAACAAAAGC	3360
3301	GCCTTAGTGAATTTGTACCTGGAAGTACAGTATATCTTGTAAATACATAAAAACAAAAGC	3360
3361	ATTTTGTGAAGGAGAAGCTAATATGATTTTTTAAAGTCTATGTTTTTAAATAATATGTAAA	3420
3361	ATTTTGTGAAGGAGAAGCTAATATGATTTTTTAAAGTCTATGTTTTTAAATAATATGTAAA	3420
3421	TTTTTCAGCTATTAGTGTATATATTTTATGGTGGGAATAAAATTTCTACTACAG	3475
3421	TTTTTCAGCTATTAGTGTATATATTTTATGGTGGGAATAAAATTTCTACTACAG	3475

RESULT 10  
PCT-US95-03718-3  
; Sequence 3, Application PC/TUS9503718  
; GENERAL INFORMATION:  
; APPLICANT: GENENTECH, INC.  
; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES THEREOF  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/03718  
; FILING DATE:

|||||  
721 CATGAATATTGGACGGACATGAAGTGTGTGCCAGAAATGAACCTGGGCAGGAATGC 780  
Db  
781 ACCAGGCTGTTACAAATAGATCTAAATCAAACTCCTCAGACCAATTCGCCACAATATTT 840  
Qy  
781 ACCAGGCTGTTACAAATAGATCTAAATCAAACTCCTCAGACCAATTCGCCACAATATTT 840  
Db  
841 CTTAAAGTAGGGAAACCTTATGATGAAGTGCAGAACTGTTCTATGTAACCATGGATTC 900  
Qy  
841 CTTAAAGTAGGGAAACCTTATGATGAAGTGCAGAACTGTTCTATGTAACCATGGATTC 900  
Db  
901 GGGCTCACCTGGGAAATAGAAAAAACAAGCACTCGAGGAGGCAACTACTTTTGAGATGAGT 960  
Qy  
901 GGGCTCACCTGGGAAATAGAAAAAACAAGCACTCGAGGAGGCAACTACTTTTGAGATGAGT 960  
Db  
961 ACCTATTCAACAAACAGAACTPATGATACGGATTCGTGTTGTTGTATCATCAGTGCCA 1020  
Qy  
961 ACCTATTCAACAAACAGAACTPATGATACGGATTCGTGTTGTTGTATCATCAGTGCCA 1020  
Db  
1021 AGAAACGACACCGGATACACTGTTGTTCTTCAAAAGCATCCAGTCAATCAGCTTTG 1080  
Qy  
1021 AGAAACGACACCGGATACACTGTTGTTCTTCAAAAGCATCCAGTCAATCAGCTTTG 1080  
Db  
1081 GTTACCATCGTAGGAAGGGATTTATAAATGCTACCAATTCAGAGGAAATTTATGAAT 1140  
Qy  
1081 GTTACCATCGTAGGAAGGGATTTATAAATGCTACCAATTCAGAGGAAATTTATGAAT 1140  
Db  
1141 GACCAATATGAAGAGTTTGTGTTTCTGTCAGGTTTAAAGGCTACCCACAATCAGATGT 1200  
Qy  
1141 GACCAATATGAAGAGTTTGTGTTTCTGTCAGGTTTAAAGGCTACCCACAATCAGATGT 1200  
Db  
1201 ACGTGGACCTTCTCTCGAAAAATCATTTCTGTGAGCAAAAAGGGTCTTGATAACGGATAC 1260  
Qy  
1201 ACGTGGACCTTCTCTCGAAAAATCATTTCTGTGAGCAAAAAGGGTCTTGATAACGGATAC 1260  
Db  
1261 AGCATATCCAAGTTTGGCAATCATGAACCAAGCAGGAGGAAATATATATTCATGCAGAA 1320  
Qy  
1261 AGCATATCCAAGTTTGGCAATCATGAACCAAGCAGGAGGAAATATATATTCATGCAGAA 1320  
Db  
1321 AATGATGATGCCAATTTACCAAAATGTTTCAGCTGAAATATAAGAAAGGAACCTCAAGTG 1380  
Qy  
1321 AATGATGATGCCAATTTACCAAAATGTTTCAGCTGAAATATAAGAAAGGAACCTCAAGTG 1380  
Db  
1381 CTCGCAAGCATCGGCAAGTCAGCAAGTCTCCAACTGCAAGAGAGATCACAGAAGA 1500  
Qy  
1381 CTCGCAAGCATCGGCAAGTCAGCAAGTCTCCAACTGCAAGAGAGATCACAGAAGA 1500  
Db  
1441 TGGACCTGGAAAGAGTGTTCAGAAAGTCTCCAACTGCAAGAGAGATCACAGAAGA 1560  
Qy  
1441 TGGACCTGGAAAGAGTGTTCAGAAAGTCTCCAACTGCAAGAGAGATCACAGAAGA 1560  
Db  
1501 GTCTGGAAATAGAAAGGCTAAACAGAAAGTGTGAGCAGTGGGTCTGAGCAGTACTCTA 1560  
Qy  
1501 GTCTGGAAATAGAAAGGCTAAACAGAAAGTGTGAGCAGTGGGTCTGAGCAGTACTCTA 1560  
Db  
1561 AACATGATGAAGCAATAAAAGGGTCTCTGTTCTCGGATGGATACCAATTCCTTGGC 1620  
Qy  
1561 AACATGATGAAGCAATAAAAGGGTCTCTGTTCTCGGATGGATACCAATTCCTTGGC 1620  
Db  
1621 ACATCTGTGAGACGATCTTTTAACTCTCAGGCCCCCTTCCCTTTCATCCAGACAAAC 1680  
Qy  
1621 ACATCTGTGAGACGATCTTTTAACTCTCAGGCCCCCTTCCCTTTCATCCAGACAAAC 1680  
Db  
1681 ATCTCATTTCTATGCAAAATGTTGTTGTTCTCTCTCTCATTTGTTTAACTCCCTGCTA 1740  
Qy  
1681 ATCTCATTTCTATGCAAAATGTTGTTGTTCTCTCTCTCTCATTTGTTTAACTCCCTGCTA 1740  
Db  
1741 ATTTGTCAAGTACAAAAAGCAATTTAGGTATGAAAGCAGCTACAGATGATCAGGTG 1800  
Qy  
1741 ATTTGTCAAGTACAAAAAGCAATTTAGGTATGAAAGCAGCTACAGATGATCAGGTG 1800  
Db  
1801 ACCGGCTCCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGAAATATGAATATGATCTC 1860  
Qy

1801 ACCGGCTCCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGAAATATGAATATGATCTC 1860  
Db  
1861 AAATGGAGCTTTCCAGAGAAATTTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920  
Qy  
1861 AAATGGAGCTTTCCAGAGAAATTTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920  
Db  
1921 GGAAGAGTGAATGAACGCAACAGCTTATAGGAATTAAGCAAAAACAGGAGTCTCAATCCAGGTT 1980  
Qy  
1921 GGAAGAGTGAATGAACGCAACAGCTTATAGGAATTAAGCAAAAACAGGAGTCTCAATCCAGGTT 1980  
Db  
1981 GCCGTCAAAATGCTGAAAGAAAGCAAGAGCTCTGAAAGAGAGGCACTCATGTCAAGAA 2040  
Qy  
1981 GCCGTCAAAATGCTGAAAGAAAGCAAGAGCTCTGAAAGAGAGGCACTCATGTCAAGAA 2040  
Db  
2041 CTCAGAGTGAACCCAGCTGGGAAGCAACAGAGATATTGTGAACCTCTCGGGGCGTGC 2100  
Qy  
2041 CTCAGAGTGAACCCAGCTGGGAAGCAACAGAGATATTGTGAACCTCTCGGGGCGTGC 2100  
Db  
2101 ACCTGTCCAGGACCAATTTTACTTCATTTTGAATACTGTTGCTATGTTGATCTTCTCAAC 2160  
Qy  
2101 ACCTGTCCAGGACCAATTTTACTTCATTTTGAATACTGTTGCTATGTTGATCTTCTCAAC 2160  
Db  
2161 TATCTAAGAAGTAAAGAGAAAAATTTTACAGGACTTTGGACAGAGATTTTCAAGGAAACAC 2220  
Qy  
2161 TATCTAAGAAGTAAAGAGAAAAATTTTACAGGACTTTGGACAGAGATTTTCAAGGAAACAC 2220  
Db  
2221 AATTTCCAGTTTTTACCCCACTTTTCCAATCAGATCCAAATTCAGCATGCTGTTCAAGA 2280  
Qy  
2221 AATTTCCAGTTTTTACCCCACTTTTCCAATCAGATCCAAATTCAGCATGCTGTTCAAGA 2280  
Db  
2281 GAAGTTCAGATACACCCGAGCTCGGATCAAAATCTCAGGGCTTCATGGGAAATTCATTTTCAAC 2340  
Qy  
2281 GAAGTTCAGATACACCCGAGCTCGGATCAAAATCTCAGGGCTTCATGGGAAATTCATTTTCAAC 2340  
Db  
2341 TCTCAAGATGAATTTGAATATGAAAAACCAAAAAGGCTTGGAAAGAGAGGAGACTTGAAT 2400  
Qy  
2341 TCTCAAGATGAATTTGAATATGAAAAACCAAAAAGGCTTGGAAAGAGAGGAGACTTGAAT 2400  
Db  
2401 GTGCTTACATTTGAAGATCTTTCTTTGCTTGGATATCAAGTTTGCCAAAGGAATGGAATTT 2460  
Qy  
2401 GTGCTTACATTTGAAGATCTTTCTTTGCTTGGATATCAAGTTTGCCAAAGGAATGGAATTT 2460  
Db  
2461 CTGGAAATTTAAGTCTGTGTTTCAAGAGACCTGGCCGCGCAGGAACTGTTGCTCACCAC 2520  
Qy  
2461 CTGGAAATTTAAGTCTGTGTTTCAAGAGACCTGGCCGCGCAGGAACTGTTGCTCACCAC 2520  
Db  
2521 GGGAAAGTGTGGAAGATATGTGATTTTGGATTTGGCTCGAGATATCATGATGATTTCCAAC 2580  
Qy  
2521 GGGAAAGTGTGGAAGATATGTGATTTTGGATTTGGCTCGAGATATCATGATGATTTCCAAC 2580  
Db  
2581 TATGTTGTTCAGGGCAATGCCCCGTCTGCTGTAAAAATGGATGGCCCCCGAAAGCCTGTTT 2640  
Qy  
2581 TATGTTGTTCAGGGCAATGCCCCGTCTGCTGTAAAAATGGATGGCCCCCGAAAGCCTGTTT 2640  
Db  
2641 GAAGGCATCTACACATTAAGAGTGAATGTCTGGTCATATGGAATATTACTGTGGGAAATC 2700  
Qy  
2641 GAAGGCATCTACACATTAAGAGTGAATGTCTGGTCATATGGAATATTACTGTGGGAAATC 2700  
Db  
2701 TTCTCAGTTGTGTAATCTTACCTGSCATTCCTGGTTGATGCTAACTTCTACAAAACATG 2760  
Qy  
2701 TTCTCAGTTGTGTAATCTTACCTGSCATTCCTGGTTGATGCTAACTTCTACAAAACATG 2760  
Db  
2761 ATTCAAAATGGATTTAAAAATGGATCAGCAATTTTATGCTACAGAAAGAAATATACATTA 2820  
Qy  
2761 ATTCAAAATGGATTTAAAAATGGATCAGCAATTTTATGCTACAGAAAGAAATATACATTA 2820  
Db  
2821 ATGCAATCCTCTGGGCTTTTGAATCAAGGAAACCGGCAATCTTCCCTTAAATTTGACTTCG 2880  
Qy  
2821 ATGCAATCCTCTGGGCTTTTGAATCAAGGAAACCGGCAATCTTCCCTTAAATTTGACTTCG 2880  
Db  
2881 TTTTGTAGGATGTACGCTGGCAGATGCAAGAGGATGATCAGAAATGTCGATGCGCGT 2940  
Qy  
2881 TTTTGTAGGATGTACGCTGGCAGATGCAAGAGGATGATCAGAAATGTCGATGCGCGT 2940  
Db

2941 GTTTCGGAATGCTCCACACCTTACCAAAACAGGCGACCTTTCCAGCAGAGAGATGATTTG 3000  
DB GTTTCGGAATGCTCCACACCTTACCAAAACAGGCGACCTTTCCAGCAGAGAGATGATTTG 3000  
3001 GGGCTACTCTCCCGAGGCTCAGTCGAGTTCGTAGAGAACTTTTACGTTTACG 3060  
DB GGGCTACTCTCCCGAGGCTCAGTCGAGTTCGTAGAGAACTTTTACGTTTACG 3060  
3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAAACAGATTAAATTTTCAT 3120  
DB ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAAACAGATTAAATTTTCAT 3120  
3121 CACTAAAAAATACTATATCAACTGTCTTCCACAGACTTTTCTCTAGAAAGCGCTCT 3180  
DB CACTAAAAAATACTATATCAACTGTCTTCCACAGACTTTTCTCTAGAAAGCGCTCT 3180  
3181 GCGTTTACTCTTTTCAAGGGACTTTTAAATCAAAATCAATCATCTGTCAAGGCG 3240  
DB GCGTTTACTCTTTTCAAGGGACTTTTAAATCAAAATCAATCATCTGTCAAGGCG 3240  
3241 GAGGAGCTGATTAATCAACTTTTATGAGCAATTCATCTGATCAAGGCGCTTCTCAGGCG 3300  
DB GAGGAGCTGATTAATCAACTTTTATGAGCAATTCATCTGATCAAGGCGCTTCTCAGGCG 3300  
3301 GCTTGAGTGAAATGTTGATCTACCTGAAGTACAGTATATTTCTGTAATACATAAAACAAAGC 3360  
DB GCTTGAGTGAAATGTTGATCTACCTGAAGTACAGTATATTTCTGTAATACATAAAACAAAGC 3360  
3361 ATTTTGCTAAGGAGAGCTAATATGATTTTAAAGTCTATGTTTAAATATATATGTAA 3420  
DB ATTTTGCTAAGGAGAGCTAATATGATTTTAAAGTCTATGTTTAAATATATATGTAA 3420  
3421 TTTTTCAGCTATTTAGTGATATATTTATGGGTGGGAATAAAATTTCTACTACAG 3475  
DB TTTTTCAGCTATTTAGTGATATATTTATGGGTGGGAATAAAATTTCTACTACAG 3475

RESULT 11

US-08-183-211-1  
; Sequence 1, Application US/08183211  
; Patent No. 5618709  
; GENERAL INFORMATION:  
; APPLICANT: Alan M. Gewirtz, Donald Small, Curt I. Civin.  
; TITLE OF INVENTION: ANTIENSE OLIGONUCLEO- TIDES  
; TITLE OF INVENTION: SPECIFIC FOR STK-1 AND METHOD FOR  
; TITLE OF INVENTION: INHIBITING EXPRESSION OF THE STK-1 PROTEIN  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: SEIDEL GONDA LAVORGNA & MONACO  
; STREET: Suite 1800, Penn Center Plaza  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: U.S.A.  
; ZIP: 19102  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 720 kb  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/183,211  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Monaco, Daniel A.  
; REGISTRATION NUMBER: 30,480  
; REFERENCE/DOCKET NUMBER: 3957-15  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-8383

; TELEFAX: (215) 568-5549  
; TELEX: No. 5618709e  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3476 nucleotides  
; TYPE: nucleic acid  
; STRANDEDNESS: single stranded  
; TOPOLOGY: linear  
; US-08-183-211-1

Query Match 99.2%; Score 3474.4; DB 1; Length 3476;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3475; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 CGAGCGGCGATCCGAGGCGTGGCGGCGCCCTGGGGGACCCCGGCGCTCCGAGGCGCATG 60  
Db 1 CGAGCGGCGATCCGAGGCGTGGCGGCGCCCTGGGGGACCCCGGCGCTCCGAGGCGCATG 60  
Qy 61 CGGCGTTGGCGCGGACCGCGGCGACCGTGGCGCTGCTGTTGTTTCTGCAATGATA 120  
Db 61 CGGCGTTGGCGCGGCGCGCGCGGCGACCGTGGCGCTGCTGTTGTTTCTGCAATGATA 120  
Qy 121 TTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTAAATCAATCAATAG 180  
Db 121 TTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTAAATCAATCAATAG 180  
Qy 181 AACAAATGATTTCACTAGTGGGGAAGTCATCATCATATCCCATGATTCAGAAATCCCGGAA 240  
Db 181 AACAAATGATTTCACTAGTGGGGAAGTCATCATCATATCCCATGATTCAGAAATCCCGGAA 240  
Qy 241 GACCTCGGCTGCGTTGAGACCCAGAGCTCAGGACAGTGTACGAAAGTGGCGGTGTG 300  
Db 241 GACCTCGGCTGCGTTGAGACCCAGAGCTCAGGACAGTGTACGAAAGTGGCGGTGTG 300  
Qy 301 GAAAGTGATGATCTGCTTCCATCACTGCAAGTCTGCTGATGCCCGGAGCAAT 360  
Db 301 GAAAGTGATGATCTGCTTCCATCACTGCAAGTCTGCTGATGCCCGGAGCAAT 360  
Qy 361 TCCTGCTCTGGGCTCTTTAAGCACAGCTCCCTGAAATGGCAGCCACATTTTGTATACAA 420  
Db 361 TCCTGCTCTGGGCTCTTTAAGCACAGCTCCCTGAAATGGCAGCCACATTTTGTATACAA 420  
Qy 421 AACAGAGGATGTTTCCATGTGCTATTTTGAATAAGACAAACCCAGCTGGAGATAC 480  
Db 421 AACAGAGGATGTTTCCATGTGCTATTTTGAATAAGACAAACCCAGCTGGAGATAC 480  
Qy 481 CTACTTTTATTCAGAGTGAAGTACCAATTAACAATATGTTTACAGTGTGATTAAGA 540  
Db 481 CTACTTTTATTCAGAGTGAAGTACCAATTAACAATATGTTTACAGTGTGATTAAGA 540  
Qy 541 AATACCTCTGTTTACACATTAAGAGACCTTACTTTAGAAAAATGAAAAACAGGACGCC 600  
Db 541 AATACCTCTGTTTACACATTAAGAGACCTTACTTTAGAAAAATGAAAAACAGGACGCC 600  
Qy 601 CTGGTCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGGAAATGGGTGCTTCCGATTCA 660  
Db 601 CTGGTCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGGAAATGGGTGCTTCCGATTCA 660  
Qy 661 CAGGGGGAAGCTGTAAAGAGAAAGTCCAGCTGTTGTTTAAAAAGAGAGAAAAAGTCTT 720  
Db 661 CAGGGGGAAGCTGTAAAGAGAAAGTCCAGCTGTTGTTTAAAAAGAGAGAAAAAGTCTT 720  
Qy 721 CATGAATTTTGGGACGACATAGGTGCTGTGCCAGAAATGAACCTGGGACGGAATGC 780  
Db 721 CATGAATTTTGGGACGACATAGGTGCTGTGCCAGAAATGAACCTGGGACGGAATGC 780  
Qy 781 ACCAGGCTGTTTCACAATAGATCTTAAATCAAACTCCTCAGACCAATGTCACCAATATTT 840  
Db 781 ACCAGGCTGTTTCACAATAGATCTTAAATCAAACTCCTCAGACCAATGTCACCAATATTT 840  
Qy 841 CTTAAAGTAGGGGAACCCCTTATGATAAGGTGCAAAAGCTGTTTCATGTGAACCATGATTC 900  
Db 841 CTTAAAGTAGGGGAACCCCTTATGATAAGGTGCAAAAGCTGTTTCATGTGAACCATGATTC 900

Qy	901	GGGCTCACCTGGGAAATTAGAAAAA	901	GGGCTCAAAATGCTGAAAGAAAAG
Db	901	GGGCTCACCTGGGAAATTAGAAAAA	Db	GGGCTCAAAATGCTGAAAGAAAAG
Qy	961	ACCTATTCAACAAACAGAACTATGAT	Qy	CTCAAGATGATGACCCAGCTGGGA
Db	961	ACCTATTCAACAAACAGAACTATGAT	Db	CTCAAGATGATGACCCAGCTGGGA
Qy	1021	AGAAACGACACCGGATACTACATG	Qy	ATCTGTCAGGACCAATTTACTTTG
Db	1021	AGAAACGACACCGGATACTACATG	Db	ATCTGTCAGGACCAATTTACTTTG
Qy	1081	GTTTACCATCTAGGAAGGATTTATA	Qy	TATCTAAGAAAGTAAAGAGAAAA
Db	1081	GTTTACCATCTAGGAAGGATTTATA	Db	TATCTAAGAAAGTAAAGAGAAAA
Qy	1141	GACCAATATGAAGATTTGTTTCTG	Qy	TAICTAAGAAAGTAAAGAGAAAA
Db	1141	GACCAATATGAAGATTTGTTTCTG	Db	TAICTAAGAAAGTAAAGAGAAAA
Qy	1201	ACGTGGACCTTCTCTCGAAAAATCA	Qy	AAITTCAGTTTTTATACCCCACT
Db	1201	ACGTGGACCTTCTCTCGAAAAATCA	Db	AAITTCAGTTTTTATACCCCACT
Qy	1261	AGCATATCCAAAGTTTGGCAATCAT	Qy	GAAGTTTCAGATACACCCGGACT
Db	1261	AGCATATCCAAAGTTTGGCAATCAT	Db	GAAGTTTCAGATACACCCGGACT
Qy	1321	AATGATGATGCCCAATTTACCAAA	Qy	GTGCTTACATTTGAAGATCTTTG
Db	1321	AATGATGATGCCCAATTTACCAAA	Db	GTGCTTACATTTGAAGATCTTTG
Qy	1381	CTCGCAGAGCATCGCAAGTCAGCG	Qy	CTGGAATTTAAGTCTGTGTTTCA
Db	1381	CTCGCAGAGCATCGCAAGTCAGCG	Db	CTGGAATTTAAGTCTGTGTTTCA
Qy	1441	TGGACCTGGAAAGAGTGTTCAGAA	Qy	GGGAAAGTGGTGAAGATATGTGA
Db	1441	TGGACCTGGAAAGAGTGTTCAGAA	Db	GGGAAAGTGGTGAAGATATGTGA
Qy	1501	GTCTGGAATAGAAAAGGCTTAAAG	Qy	TATGTTGTGTCAGGGGCAATG
Db	1501	GTCTGGAATAGAAAAGGCTTAAAG	Db	TATGTTGTGTCAGGGGCAATG
Qy	1561	AAATGATGATGAAAGGCTTAAAG	Qy	GAAGGTCATACACCAATTAAG
Db	1561	AAATGATGATGAAAGGCTTAAAG	Db	GAAGGTCATACACCAATTAAG
Qy	1621	ACATCTTTGTGAGACGATCCTTTT	Qy	TTCTCAGTTGGTGTGAATCTTAC
Db	1621	ACATCTTTGTGAGACGATCCTTTT	Db	TTCTCAGTTGGTGTGAATCTTAC
Qy	1681	ATCTCATCTATGCAACAAATTTG	Qy	ATTCAAATGGAATTTAAATGGA
Db	1681	ATCTCATCTATGCAACAAATTTG	Db	ATTCAAATGGAATTTAAATGGA
Qy	1741	ATTTGTCAACAAGTACAAAAGCA	Qy	ATGCAATCTCTCTGGGCTTTTG
Db	1741	ATTTGTCAACAAGTACAAAAGCA	Db	ATGCAATCTCTCTGGGCTTTTG
Qy	1801	ACCGGCTCCTCAGATATGAGTACT	Qy	TTTTCAGATGTCAGTGGCAGAT
Db	1801	ACCGGCTCCTCAGATATGAGTACT	Db	TTTTCAGATGTCAGTGGCAGAT
Qy	1861	AAATGGGATTTCCAGAGAAAA	Qy	GTTCGGAATGTCCTCACCTAC
Db	1861	AAATGGGATTTCCAGAGAAAA	Db	GTTCGGAATGTCCTCACCTAC
Qy	1921	GGAAAGTGATGAAAGGCTTATG	Qy	GGGCTACTCTCTCCGCGAGGCT
Db	1921	GGAAAGTGATGAAAGGCTTATG	Db	GGGCTACTCTCTCCGCGAGGCT
Qy	901	GGGCTCACCTGGGAAATTAGAAAAA	Qy	GGGCTCAAAATGCTGAAAGAAAAG
Db	901	GGGCTCACCTGGGAAATTAGAAAAA	Db	GGGCTCAAAATGCTGAAAGAAAAG
Qy	961	ACCTATTCAACAAACAGAACTATGAT	Qy	CTCAAGATGATGACCCAGCTGGGA
Db	961	ACCTATTCAACAAACAGAACTATGAT	Db	CTCAAGATGATGACCCAGCTGGGA
Qy	1021	AGAAACGACACCGGATACTACATG	Qy	ATCTGTCAGGACCAATTTACTTTG
Db	1021	AGAAACGACACCGGATACTACATG	Db	ATCTGTCAGGACCAATTTACTTTG
Qy	1081	GTTTACCATCTAGGAAGGATTTATA	Qy	TATCTAAGAAAGTAAAGAGAAAA
Db	1081	GTTTACCATCTAGGAAGGATTTATA	Db	TATCTAAGAAAGTAAAGAGAAAA
Qy	1141	GACCAATATGAAGATTTGTTTCTG	Qy	TAICTAAGAAAGTAAAGAGAAAA
Db	1141	GACCAATATGAAGATTTGTTTCTG	Db	TAICTAAGAAAGTAAAGAGAAAA
Qy	1201	ACGTGGACCTTCTCTCGAAAAATCA	Qy	AAITTCAGTTTTTATACCCCACT
Db	1201	ACGTGGACCTTCTCTCGAAAAATCA	Db	AAITTCAGTTTTTATACCCCACT
Qy	1261	AGCATATCCAAAGTTTGGCAATCAT	Qy	GAAGTTTCAGATACACCCGGACT
Db	1261	AGCATATCCAAAGTTTGGCAATCAT	Db	GAAGTTTCAGATACACCCGGACT
Qy	1321	AATGATGATGCCCAATTTACCAAA	Qy	GTGCTTACATTTGAAGATCTTTG
Db	1321	AATGATGATGCCCAATTTACCAAA	Db	GTGCTTACATTTGAAGATCTTTG
Qy	1381	CTCGCAGAGCATCGCAAGTCAGCG	Qy	CTGGAATTTAAGTCTGTGTTTCA
Db	1381	CTCGCAGAGCATCGCAAGTCAGCG	Db	CTGGAATTTAAGTCTGTGTTTCA
Qy	1441	TGGACCTGGAAAGAGTGTTCAGAA	Qy	GGGAAAGTGGTGAAGATATGTGA
Db	1441	TGGACCTGGAAAGAGTGTTCAGAA	Db	GGGAAAGTGGTGAAGATATGTGA
Qy	1501	GTCTGGAATAGAAAAGGCTTAAAG	Qy	TATGTTGTGTCAGGGGCAATG
Db	1501	GTCTGGAATAGAAAAGGCTTAAAG	Db	TATGTTGTGTCAGGGGCAATG
Qy	1561	AAATGATGATGAAAGGCTTAAAG	Qy	GAAGGTCATACACCAATTAAG
Db	1561	AAATGATGATGAAAGGCTTAAAG	Db	GAAGGTCATACACCAATTAAG
Qy	1621	ACATCTTTGTGAGACGATCCTTTT	Qy	TTCTCAGTTGGTGTGAATCTTAC
Db	1621	ACATCTTTGTGAGACGATCCTTTT	Db	TTCTCAGTTGGTGTGAATCTTAC
Qy	1681	ATCTCATCTATGCAACAAATTTG	Qy	ATTCAAATGGAATTTAAATGGA
Db	1681	ATCTCATCTATGCAACAAATTTG	Db	ATTCAAATGGAATTTAAATGGA
Qy	1741	ATTTGTCAACAAGTACAAAAGCA	Qy	ATGCAATCTCTCTGGGCTTTTG
Db	1741	ATTTGTCAACAAGTACAAAAGCA	Db	ATGCAATCTCTCTGGGCTTTTG
Qy	1801	ACCGGCTCCTCAGATATGAGTACT	Qy	TTTTCAGATGTCAGTGGCAGAT
Db	1801	ACCGGCTCCTCAGATATGAGTACT	Db	TTTTCAGATGTCAGTGGCAGAT
Qy	1861	AAATGGGATTTCCAGAGAAAA	Qy	GTTCGGAATGTCCTCACCTAC
Db	1861	AAATGGGATTTCCAGAGAAAA	Db	GTTCGGAATGTCCTCACCTAC
Qy	1921	GGAAAGTGATGAAAGGCTTATG	Qy	GGGCTACTCTCTCCGCGAGGCT
Db	1921	GGAAAGTGATGAAAGGCTTATG	Db	GGGCTACTCTCTCCGCGAGGCT

Db 3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACAAAACAGATTATTTTCAT 3120  
QY 3121 CACTAAAGAAATCTATTATCAACTCTGCTTCCACAGACTTTTCTCTAGAACCGCTCT 3180  
Db 3121 CACTAAAGAAATCTATTATCAACTCTGCTTCCACAGACTTTTCTCTAGAACCGCTCT 3180  
QY 3181 GCCTTTACTCTGTTTCAAGGACCTTTTGTAAATCAAAATCAATCTCTGTACAAAGCAG 3240  
Db 3181 GCCTTTACTCTGTTTCAAGGACCTTTTGTAAATCAAAATCAATCTCTGTACAAAGCAG 3240  
QY 3241 GAGGAGCTGATAATGAATCTTTATTTGAGCATTTGATCTGATCCAAAGGCTTTCTCAGGCGG 3300  
Db 3241 GAGGAGCTGATAATGAATCTTTATTTGAGCATTTGATCTGATCCAAAGGCTTTCTCAGGCGG 3300  
QY 3301 GCCTTTGAGTGAATTTGTGACCTGAACTGACAGTATATTTCTGTAAATACATAAAACAAAAGC 3360  
Db 3301 GCCTTTGAGTGAATTTGTGACCTGAACTGACAGTATATTTCTGTAAATACATAAAACAAAAGC 3360  
QY 3361 ATTTTGTAAAGGAGAACTAATATGATTTTAAATCTATGTTTAAATATATGTA 3420  
Db 3361 ATTTTGTAAAGGAGAACTAATATGATTTTAAATCTATGTTTAAATATATGTA 3420  
QY 3421 TTTTTCAGCTATTAGTGATATATTTTATGAGTGGGATATAAAATTTCTACTACAGA 3476  
Db 3421 TTTTTCAGCTATTAGTGATATATTTTATGAGTGGGATATAAAATTTCTACTACAGA 3476

## RESULT 12

PC-US95-00176A-1  
; Sequence 1, Application PC/US9500176A  
; GENERAL INFORMATION:  
; APPLICANT: Alan M. Gewirtz, Donald Small, Curt I. Civin.  
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDES  
; TITLE OF INVENTION: SPECIFIC FOR SYK-1 AND METHOD FOR INHIBITING EXPRESSION OF THE  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEIDEL GONDA LAVORNA & MONACO  
; STREET: Suite 1800, Penn Center Plaza  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: U.S.A.  
; ZIP: 19102  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 720 KB  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/00176A  
; FILING DATE: 6 January 1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/183,211  
; FILING DATE: 14 January 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Monaco, Daniel A.  
; REGISTRATION NUMBER: 30,480  
; REFERENCE/DOCKET NUMBER: 3957-14 PC  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-8383  
; TELEFAX: (215) 568-5549  
; TELEX: None  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3476 nucleotides  
; TYPE: nucleic acid  
; STRANDEDNESS: single stranded  
; TOPOLOGY: linear  
PC-US95-00176A-1

Query Match 99.2%; Score 3474.4; DB 5; Length 3476;  
Best Local Similarity 100.0%; Pred. No. 0;

Matches 3475; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CGAGGCGGCATCCGAGGCTGGGCGGCGCTGGGGAGCCCGGGCTCGGAGGCCCATG 60  
Db 1 CGAGGCGGCATCCGAGGCTGGGCGGCGCTGGGGAGCCCGGGCTCGGAGGCCCATG 60  
QY 61 CGGCGGTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120  
Db 61 CGGCGGTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120  
QY 121 TTTGGGACTATTACAAATCAAGATCTGCTGTGTGATCAAGTGTGTTTAAATCAATCAAG 180  
Db 121 TTTGGGACTATTACAAATCAAGATCTGCTGTGTGATCAAGTGTGTTTAAATCAATCAAG 180  
QY 181 AACAAATGATTCATCAGTGGGAAATCATCATATATCCATGTGTATCAAGATCCCGGAA 240  
Db 181 AACAAATGATTCATCAGTGGGAAATCATCATATATCCATGTGTATCAAGATCCCGGAA 240  
QY 241 GACCTCGGCTGTGGTGTGAGACCCAGAGCTCAGGAGCAGGTGTACGAGCTGCGCTGTG 300  
Db 241 GACCTCGGCTGTGGTGTGAGACCCAGAGCTCAGGAGCAGGTGTACGAGCTGCGCTGTG 300  
QY 301 GAAGTGATGTATCTGCTTCCATCACAAGTGTGCTGTGCTGATGCCCGAGGAAAT 360  
Db 301 GAAGTGATGTATCTGCTTCCATCACAAGTGTGCTGTGCTGATGCCCGAGGAAAT 360  
QY 361 TCCTGTCTCTGGGCTTTTAAAGCAGAGCTCCCTGAAATTCGCGAGCACAATTTGATTA 420  
Db 361 TCCTGTCTCTGGGCTTTTAAAGCAGAGCTCCCTGAAATTCGCGAGCACAATTTGATTA 420  
QY 421 AACAGAGGAGTGTTCATGCTTCCATGCTTGTGAAATGACAGAAACCCAGAGCTGGAA 480  
Db 421 AACAGAGGAGTGTTCATGCTTCCATGCTTGTGAAATGACAGAAACCCAGAGCTGGAA 480  
QY 481 CTACTTTTATTCAGAGTGAAGTACCAATTTACAAATTTGTTTACAGTGTAGTATAAGA 540  
Db 481 CTACTTTTATTCAGAGTGAAGTACCAATTTACAAATTTGTTTACAGTGTAGTATAAGA 540  
QY 541 AATACCTCTGCTTTACATTAAGAGAGCTTTTACATTTAGAAATTTGAAACCCAGAGCC 600  
Db 541 AATACCTCTGCTTTACATTAAGAGAGCTTTTACATTTAGAAATTTGAAACCCAGAGCC 600  
QY 601 CTGCTCTGCATATCTGAGAGCGTTCAGAGCGGATCGTGGAAATGGGTGCTTTGCCATCA 660  
Db 601 CTGCTCTGCATATCTGAGAGCGTTCAGAGCGGATCGTGGAAATGGGTGCTTTGCCATCA 660  
QY 661 CAGGGGAAAGCTGTAAAGAGAGAGTCCAGCTGTGTTTAAAGAGAGAGAGAGAGTGT 720  
Db 661 CAGGGGAAAGCTGTAAAGAGAGAGTCCAGCTGTGTTTAAAGAGAGAGAGAGAGTGT 720  
QY 721 CATGAAATTTATTTGGGACGGACATTAAGGTGCTGTGCCAGAAATGAATCGGCGAGGAA 780  
Db 721 CATGAAATTTATTTGGGACGGACATTAAGGTGCTGTGCCAGAAATGAATCGGCGAGGAA 780  
QY 781 ACCAGGCTGTTTCAAAATAGATCTTAAATCAAACTCTCTAGACACATTCGCAATATTT 840  
Db 781 ACCAGGCTGTTTCAAAATAGATCTTAAATCAAACTCTCTAGACACATTCGCAATATTT 840  
QY 841 CTTAAAGTAGGGGAAACCTTTATGATAGGTGCAAGCTGTTCATGTGAACCATGATTC 900  
Db 841 CTTAAAGTAGGGGAAACCTTTATGATAGGTGCAAGCTGTTCATGTGAACCATGATTC 900  
QY 901 GGGCTCACCTGGGAAATTAAGAAACAAAGCACTCGAGGAGGCACTACTTTGAGATGAGT 960  
Db 901 GGGCTCACCTGGGAAATTAAGAAACAAAGCACTCGAGGAGGCACTACTTTGAGATGAGT 960  
QY 961 ACCTATTCAACAAACAGAACTATGATACGATTCCTTTGCTTTGTTTATCATCATGCGCA 1020  
Db 961 ACCTATTCAACAAACAGAACTATGATACGATTCCTTTGCTTTGTTTATCATCATGCGCA 1020  
QY 1021 AGAAACGACACCGGATACACTACCTTTGCTTTCAAGAGCATCCAGTCAATCAGCTTTG 1080  
Db 1021 AGAAACGACACCGGATACACTACCTTTGCTTTCAAGAGCATCCAGTCAATCAGCTTTG 1080



QY	1081	GTTACCATCGTAGGAAGGATTTATATAATCTACCAATTCAGTGAAGATTATGAATT	1140
Db	1081		
QY	1141	GACCAATATGAAGATTGTTGTTTCTGTGAGGTTTAAAGCCTACCCACAAATCAGATGT	1200
Db	1141		
QY	1201	ACGTGGACCTTCTCTCGAAATCATTTCTTGTGAGCAAAAGGCTTTGATPAACGGATAC	1260
Db	1201		
QY	1261	AGCATATCCAAATTTTCCAAATCATTAAGCACCAGCCAGGAGATATATATCCATGCGAA	1320
Db	1261		
QY	1321	AATGATGATGCCCAATTTACCAAAATGTTCAAGCTGAATATAAGAGGAACCTCAAGTG	1380
Db	1321		
QY	1381	CTCGCAAGACATCGGCAAGTCAAGCGTCTGTTTCTCGATGATACCCATTACCATCT	1440
Db	1381		
QY	1441	TGGACCTGGAAGAGTGTTCAGACAAAGTCTCCAACTGCCACAGAGAGATCACAAGGA	1500
Db	1441		
QY	1501	GTCGGAATAGAAAGGTTAAAGAAAGTGTGAGACAGTGGGTGTCGAGCAGTACTCTA	1560
Db	1501		
QY	1561	AACATGAGTGAAGCCATAAAAGGTTCTGTGTCAGTGTGTCATACAAATCCCTTGGC	1620
Db	1561		
QY	1621	ACATCTGTGAGACGATCCTTTTAACTCTCAGGCGCCCTTCCCTTTTCATCCAGACAAC	1680
Db	1621		
QY	1681	ATCTCAATCTATGCAAAATTTGTTGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1740
Db	1681		
QY	1741	ATTTGTCAAAATGACAAAGCAATTTAGATATGAAGCCAGCTACAGATGCTACAGGTG	1800
Db	1741		
QY	1801	ACGGCTCTCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGAATATGAATGATCTC	1860
Db	1801		
QY	1861	AAATGGAGTTTCCAGAGAAATTTAGAGTTTGGGAAGTACTAGGATCAGGTGCTTTT	1920
Db	1861		
QY	1921	GGAAAGTGTATGAACGCAACAGCTTATGGAATTTAGCAAAACAGAGTCTCAATCCAGTT	1980
Db	1921		
QY	1981	GCGTCAAAATGCTGAAAGAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAA	2040
Db	1981		
QY	2041	CTCAGATGATGACCCAGCTGGGAAGCCAGAGATATGTGAACTCTGGGGGGGTGC	2100
Db	2041		
QY	2101	ACACTGTGAGGACCAATTTTACTTGTATGTTTGTGATCTGTTGATGATCTCTCAAC	2160
Db	2101		

QY	2161	TATCTAAGAGTAAAGAGAAATTTTTCACAGGACTTGGACAGAGATTTTCAAGAACAC	2220
Db	2161		
QY	2221	AAATTTCAAGTTTTCACCCACCTTTCCAAATCAGATCCAAATTCAGCATGCTGTTCAAGA	2280
Db	2221		
QY	2281	GAACTTCAGATACACCCCGACTCGATCAAATCTCAGGCTTCATGGGAATTCATTTTCA	2340
Db	2281		
QY	2341	TCTCAAGATGAAATTTGAATATGAATATGAATATGAATATGAATATGAATATGAAT	2400
Db	2341		
QY	2401	GTGCTTACATTTGAAGATCTTCTTTGTTTGCATATCAAGTTGCCAAAGGAATGAAATTT	2460
Db	2401		
QY	2461	CTGGAATTTAAGTCTGTGTTTCACAGAGACTGCGCCGACAGAACTGTTGTCAACCAC	2520
Db	2461		
QY	2521	GGGAACTGCTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGAGTGAATTCAC	2580
Db	2521		
QY	2581	TATGTTCTCAGGGCAATGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2640
Db	2581		
QY	2641	GAAGGCATCTACACCATTAAGAGTGTCTGTGCTCATATGGAAATTTACTCTGGGAAATC	2700
Db	2641		
QY	2701	TTCTCACTCTGTTGAATCTTACCTTGGCAATCCGGTTGATGCTAACTTACAAAATCG	2760
Db	2701		
QY	2761	ATTCAAATCGATTTAAATTTGAATGATGATGATGATGATGATGATGATGATGATGATG	2820
Db	2761		
QY	2821	ATGCAATCTCTGCTGGGCTTTTGAATCAAGGAAACGGCCATCTTCCCTAAATTTGAC	2880
Db	2821		
QY	2881	TTTTTAGGATGTCAGCTGGCAGATGCAAGAAAGCGATGTATCAGAAATGTGGATGGCG	2940
Db	2881		
QY	2941	GTTTCGGAATGTCCTCACACTTACCAAAACAGGGGACCTTTTCAGCAGAGATGGATTG	3000
Db	2941		
QY	3001	GGCTACTCTCTCGCAGGCTCAGGTGGAAGATTCTGAGAGGACAAATTTGATTTAAGG	3060
Db	3001		
QY	3061	ACTTCATCCCTCCACCTATCCCTTAAACAGGCTGTAGATTACAAAACAAAGATTATTT	3120
Db	3061		
QY	3121	CACATAAAGAAATCTATTATCACTGCTGCTTCAACAGATTTTCTCTAGAACCGCTCT	3180
Db	3121		
QY	3181	GGGTTTACTCTGTTTCAAGGGAATTTGTAATCAATCATCTGTGTCAGAGGAG	3240
Db	3181		
QY	3241	GAGGAGCTGATAATGAATTTTATTTGGAGCATTTGATCTGCATCCAAAGGCTTCTCAGG	3300
Db	3241		



Db 3241 GAGGAGCTGATAATGAATTTATTTGAGCAATGATCTGCATCAAAGGCCCTTCTCAGGCGG 3300  
 QY 3301 GCTTCAGTGAAATTTGTCCTGAGTACAGTATATTTCTTTGTAATAATACATAAAACAAAGC 3360  
 Db 3301 GCTTCAGTGAAATTTGTCCTGAGTACAGTATATTTCTTTGTAATAATACATAAAACAAAGC 3360  
 QY 3361 ATTTTGTAAAGGAGAGCTAAATATGATTTTTTTAAAGTCTAATGTTTTTAAATAATATGTA 3420  
 Db 3361 ATTTTGTAAAGGAGAGCTAAATATGATTTTTTTAAAGTCTAATGTTTTTAAATAATATGTA 3420  
 QY 3421 TTTTTCAGCTATTAGTATATATTTTATGCTGGGTAATAAATTTTCTACTACAGA 3476  
 Db 3421 TTTTTCAGCTATTAGTATATATTTTATGCTGGGTAATAAATTTTCTACTACAGA 3476

RESULT 13  
 US-08-222-616-22  
 ; Sequence 22, Application US/08222616  
 ; Patent No. 5635177  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bennett, Brian D.  
 ; APPLICANT: Goeddel, David  
 ; APPLICANT: Lee, James M.  
 ; APPLICANT: Matthews, William  
 ; APPLICANT: Tsai, Siao Ping  
 ; APPLICANT: Wood, William I.  
 ; TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST  
 ; TITLE OF INVENTION: ANTIBODIES  
 ; NUMBER OF SEQUENCES: 42  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Genentech, Inc.  
 ; STREET: 460 Point San Bruno Blvd  
 ; CITY: South San Francisco  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 94080  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patin (Genentech)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/222.616  
 ; FILING DATE: 4-APR-1994  
 ; CLASSIFICATION: 530  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US93/00586  
 ; FILING DATE: 22-JAN-1993  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 07/826935  
 ; FILING DATE: 22-JAN-1992  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Lee, Wendy M.  
 ; REGISTRATION NUMBER:  
 ; REFERENCE/DOCKET NUMBER: 821P2  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 415/225-1994  
 ; TELEFAX: 415/952-9881  
 ; TELEX: 910/371-7168  
 ; INFORMATION FOR SEQ ID NO: 22:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 3120 bases  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; US-08-222-616-22

Query Match 88.1%; Score 3083.2; DB 1; Length 3120;  
 Best,Local Similarity 99.3%; Pred.No. 0;  
 Matches 3097; Conservative 0; Mismatches 23; Indels 0; Gaps 0;  
 QY 53 ATGCGGCGTGGCGCGGACCGGCGCACTGCTGCTGTTGTTTCTGCAATG 117

Db 1 ATGAGAGGCTGGCGGCGGACCGGCGCGAGCTGCGCGTCTGTTGTTTCTGCAATG 60  
 QY 118 ATATTTGGGACTATTAACAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATCAT 177  
 Db 61 ATATTTGGGACTATTAACAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATCAT 120  
 QY 178 AAGAACATGATTCATCAGTGGGGAAGTCAATCATCATATATCCCATGATCAGATCCCCG 237  
 Db 121 AAGAACATGATTCATCAGTGGGGAAGTCAATCATCATATATCCCATGATCAGATCCCCG 180  
 QY 238 GAAGACCTCGGCTGTGCGTTGAGACCCAGAGCTCAGGCGACAGCTGTACGAAGCTGCGCGT 297  
 Db 181 GAAGACCTCGGCTGTGCGTTGAGACCCAGAGCTCAGGCGACAGCTGTACGAAGCTGCGCGT 240  
 QY 298 GTGGAAGTGGATGTATCTGCTTCCATCACA CTGCAAGTGTGCTGATGCGGCGGGAAC 357  
 Db 241 GTGGAAGTGGATGTATCTGCTTCCATCACA CTGCAAGTGTGCTGATGCGGCGGGAAC 300  
 QY 358 ATTTCTGTCTCTGGGTCTTTAAGGACAGCTCCCTGAATGCCAGGCACATTTTGTATTA 417  
 Db 301 ATTTCTGTCTCTGGGTCTTTAAGGACAGCTCCCTGAATGCCAGGCACATTTTGTATTA 360  
 QY 418 CAAACAGAGGAGTGTGTTTCCATGTCATTTTGAATAACAGACAGAAACCCAGCTGGAGAA 477  
 Db 361 CAAACAGAGGAGTGTGTTTCCATGTCATTTTGAATAACAGACAGAAACCCAGCTGGAGAA 420  
 QY 478 TACCTACTTTTATTCAGAGTGAAGCTACCAATTAACAATAATTTGTTTACAGTGAGTATA 537  
 Db 421 TACCTACTTTTATTCAGAGTGAAGCTACCAATTAACAATAATTTGTTTACAGTGAGTATA 480  
 QY 538 AGAATACCTGCTTTTACATTAAGAAGACCTTACTTTAGAAATAATGGAACACAGGAC 597  
 Db 481 AGAATACCTGCTTTTACATTAAGAAGACCTTACTTTAGAAATAATGGAACACAGGAC 540  
 QY 598 GCGCTGTCTGCATATCTCAGAGCGTTCCAGAGCCGATCGTGAAATGGGTCCTTTGCGAT 657  
 Db 541 GCGCTGTCTGCATATCTCAGAGCGTTCCAGAGCCGATCGTGAAATGGGTCCTTTGCGAT 600  
 QY 658 TCACAGGCGGAAAGCTGTAAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAAAGTG 717  
 Db 601 TCACAGGCGGAAAGCTGTAAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAAAGTG 660  
 QY 718 CTTTCATGAATTTATTTGGGACGACATAAGGTGCTGTGCCAGAAATGAATGGGCGAGGAA 777  
 Db 661 CTTTCATGAATTTATTTGGGACGACATAAGGTGCTGTGCCAGAAATGAATGGGCGAGGAA 720  
 QY 778 TGCAACGAGGCTGTTTACAAATAGATCTAATCAAACTCCTCAGACCAATTTGCCACAATTA 837  
 Db 721 TGCAACGAGGCTGTTTACAAATAGATCTAATCAAACTCCTCAGACCAATTTGCCACAATTA 780  
 QY 838 TTTCTTAAAGTAGGGGAACCCCTTATGATAGGTGCAAGCTGTTTCAATGTGAACCATGGA 897  
 Db 781 TTTCTTAAAGTAGGGGAACCCCTTATGATAGGTGCAAGCTGTTTCAATGTGAACCATGGA 840  
 QY 898 TTCGGGCTCACTTGGGAAATTTAGAAAAA CAAAGACTCGAGAGGGCAACTACTTTGAGATG 957  
 Db 841 TTCGGGCTCACTTGGGAAATTTAGAAAAA CAAAGACTCGAGAGGGCAACTACTTTGAGATG 900  
 QY 958 AGTACCTATTCAACAAACAGAACTATGATACGGATCTGTTGCTTTTGTATCATCATGATG 1017  
 Db 901 AGTACCTATTCAACAAACAGAACTATGATACGGATCTGTTGCTTTTGTATCATCATGATG 960  
 QY 1018 GCAAGAAACGACACCGGATACACTTGTCTCTTCAAGGACTCCAGTCAATCAGCT 1077  
 Db 961 GCAAGAAACGACACCGGATACACTTGTCTCTTCAAGGACTCCAGTCAATCAGCT 1020  
 QY 1078 TTGTTTACCATCGTAGGAAGGGATTATTAATATGCTACCAATTCAGTGAAGATTATGAA 1137  
 Db 1021 TTGTTTACCATCGTAGGAAGGGATTATTAATATGCTACCAATTCAGTGAAGATTATGAA 1080  
 QY 1138 ATTGACCAATATGAAGATTTGTTTCTGTCAGGTTTAAAGCTTACCACCAATCAGA 1197

1081 ATTGACCAATATGAAGAGTTTGTGTTTCTGTGAGGTTTAAAGCCTACCCACAATCAGA 1140  
1198 TGTACGTGGACCTTCTCTCGAATATCAATTTCCCTTGTGAGCAAAAGGCTCTGTATACGGA 1257  
1141 TGTACGTGGACCTTCTCTCGAATATCAATTTCCCTTGTGAGCAAAAGGCTCTGTATACGGA 1200  
1258 TACAGCATATCCAAAGTTTGTCAATCAATAAGCACAGCCAGGAGAAATATATTTCCATGCA 1317  
1201 TACAGCATATCCAAAGTTTGTCAATCAATAAGCACAGCCAGGAGAAATATATTTCCATGCA 1260  
1318 GAAATATGATGCCCAATTTACAAATATGTTCAAGCTGTAATATAAGAGAAACCTCAA 1377  
1261 GAAATATGATGCCCAATTTACAAATATGTTCAAGCTGTAATATAAGAGAAACCTCAA 1320  
1378 GTGCTCGCAGAGCATCGGCAAGTCTGAGCAAGTCTCCAACTGCAAGAGAGATCACAGAA 1437  
1321 GTGCTCGCAGAGCATCGGCAAGTCTGAGCAAGTCTCCAACTGCAAGAGAGATCACAGAA 1380  
1438 TCTTGGACCTGGAAAGTGTTCAGACAAAGTCTCCAACTGCAAGAGAGATCACAGAA 1497  
1381 TCTTGGACCTGGAAAGTGTTCAGACAAAGTCTCCAACTGCAAGAGAGATCACAGAA 1440  
1498 GGAGTCTGGAATAGAAAGGCTAAACAGAAAGTGTTCGCAAGTGGTGTGCGAGCACTACT 1557  
1441 GGAGTCTGGAATAGAAAGGCTAAACAGAAAGTGTTCGCAAGTGGTGTGCGAGCACTACT 1500  
1558 CTAACATGAGTGAAGCCATAAAGGTTCTCGTCAAGTGTGTCGATACAAATTTCCCTT 1617  
1501 CTAACATGAGTGAAGCCATAAAGGTTCTCGTCAAGTGTGTCGATACAAATTTCCCTT 1560  
1618 GGCACATCTTGTGAGACGATCTTTTAACTCTCCAGGCCCTTCCCTTTCATCCAGAC 1677  
1561 GGCACATCTTGTGAGACGATCTTTTAACTCTCCAGGCCCTTCCCTTTCATCCAGAC 1620  
1678 AACATCTCAATTCATGCAACAAATTTGTGTTGTCTCTCTTCAATGTCGTTTAAACCTTG 1737  
1621 AACATCTCAATTCATGCAACAAATTTGTGTTGTCTCTCTTCAATGTCGTTTAAACCTTG 1680  
1738 CTAATTTGTACAGTACAAAGGCAATTTAGTATGAAGCCAGCTACAGATGGTACAG 1797  
1681 CTAATTTGTACAGTACAAAGGCAATTTAGTATGAAGCCAGCTACAGATGGTACAG 1740  
1798 GTGACGGCTCTCAGATATGATGATCTTCAAGTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 1857  
1741 GTGACGGCTCTCAGATATGATGATCTTCAAGTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 1800  
1858 CTCAAATGGGAGTTTCCAGAGAGAAATTTAGAGTTTGGGAAGTACTAGGATCAGGTGCT 1917  
1801 GTCAATGGGAGTTTCCAGAGAGAAATTTAGAGTTTGGGAAGTACTAGGATCAGGTGCT 1860  
1918 TTTGAAAGGTGATGAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAG 1977  
1861 TTTGAAAGGTGATGAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAG 1920  
1978 GTTGGCGTCAAAATGCTGAAGAAAGCAAGCAGCTCTGAAGAGAGGCACTCATGTCA 2037  
1921 GTTACCGTCAAAATGCTGAAGAAAGCAAGCAGCTCTGAAGAGAGGCACTCATGTCA 1980  
2038 GAACTCAAGATGATGACCGCTGGGAAGCCAGAGAAATTTGTGAACTGCTGGGGGGG 2097  
1981 GAACTCAAGATGATGACCGCTGGGAAGCCAGAGAAATTTGTGAACTGCTGGGGGGG 2040  
2098 TGCACTGTCTAGGACCAATTTACTGTTTGTGAACTGTTGCTATGTTGCTATGTTGCTATGTTGCT 2157  
2041 TGCACTGTCTAGGACCAATTTACTGTTTGTGAACTGTTGCTATGTTGCTATGTTGCTATGTTGCT 2100  
2158 AACTATCTAAGAGTAAAGAGAAATTTCAAGGACTTTGACAGAGATTTTCAAGGAA 2217  
2101 AACTATCTAAGAGTAAAGAGAAATTTCAAGGACTTTGACAGAGATTTTCAAGGAA 2160  
2218 CACAATTTCAATTTTACCCCACTTTCCAACTACATCCAAATTTCCAGCATGCTGGTTCA 2277  
2161 CACAATTTCAATTTTACCCCACTTTCCAACTACATCCAAATTTCCAGCATGCTGGTTCA 2220

2278 AGAGAGTTTCAGATACACCCGGACTCGGATCAATCTCAGGGCTTCATGGGAATTCATTT 2337  
2221 AGAGAGTTTCAGATACACCCGGACTCGGATCAATCTCAGGGCTTCATGGGAATTCATTT 2280  
2338 CACTCTGAAGATCAAAATTTGAATATGAAAAACAAAAAGGCTGGAAGAGAGAGGACTTG 2397  
2281 CACTCTGAAGATCAAAATTTGAATATGAAAAACAAAAAGGCTGGAAGAGAGAGGACTTG 2340  
2398 AATGTGCTTACATTTGAAGATCTTTTGTGTTGCTATCATCAAGTTGCCAAAGGAATGAA 2457  
2341 AATGTGCTTACATTTGAAGATCTTTTGTGTTGCTATCATCAAGTTGCCAAAGGAATGAA 2400  
2458 TTTCTGGAATTTAAGTCTGTGTTTCAGAGACCTGCGCCGAGGAACGTCTGTCTAC 2517  
2401 TTTCTGGAATTTAAGTCTGTGTTTCAGAGACCTGCGCCGAGGAACGTCTGTCTAC 2460  
2518 CACGGAAAGTGTGAAGATATGTGACTTTTGGATTGGCTCGAGATATCATCAGAGTATTC 2577  
2461 CACGGAAAGTGTGAAGATATGTGACTTTTGGATTGGCTCGAGATATCATCAGAGTATTC 2520  
2578 AACTATGTTTGTACAGGGCAATGCGCGTCTGCTGTAAATGGATGGCCCGCAAGGCTG 2637  
2521 AACTATGTTTGTACAGGGCAATGCGCGTCTGCTGTAAATGGATGGCCCGCAAGGCTG 2580  
2638 TTTGAAGGCATCTACACCAATTAAGAGTGTCTGTGTCATATGGAATATTAATCTGTGGAA 2697  
2581 TTTGAAGGCATCTACACCAATTAAGAGTGTCTGTGTCATATGGAATATTAATCTGTGGAA 2640  
2698 ATCTTCTCACTTGGTGTGATCTTCACTCCGCGCAATCCCGTGTGATGCTACTTCTACAA 2757  
2641 ATCTTCTCACTTGGTGTGATCTTCACTCCGCGCAATCCCGTGTGATGCTACTTCTACAA 2700  
2758 CTGATTTCAAAATGGATTTTAAATGGATCAGCCATTTTATGCTACAGAAAGAAATATACAT 2817  
2701 CTGATTTCAAAATGGATTTTAAATGGATCAGCCATTTTATGCTACAGAAAGAAATATACAT 2760  
2818 ATAAATGCAATCTCTGCTGGGCTTTTGAATCAAGGAAACGGCCATCTTCCCTTAAATTTGACT 2877  
2761 ATAAATGCAATCTCTGCTGGGCTTTTGAATCAAGGAAACGGCCATCTTCCCTTAAATTTGACT 2820  
2878 TCGTTTTTATGAGTGTACAGCTGGCAGATGCAAGAGAGAGGATGTATCAGAATGTGGATGGC 2937  
2821 TCGTTTTTATGAGTGTACAGCTGGCAGATGCAAGAGAGAGGATGTATCAGAATGTGGATGGC 2880  
2938 CGTGTTCGGAATGTCTCTACACCTACCAAAACAGGCGACTTTTACAGCAGAGAGATGGAT 2997  
2881 CGTGTTCGGAATGTCTCTACACCTACCAAAACAGGCGACTTTTACAGCAGAGAGATGGAT 2940  
2998 TTGGGGCTACTCTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGGAACAAATTTAGTTTTA 3057  
2941 TTGGGGCTACTCTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGGAACAAATTTAGTTTTA 3000  
3058 AGGACTTCACTCCCTCACCTATCCCTAAACAGGCTGTAGATTAACAAACAGGATTAATTT 3117  
3001 AGGACTTCACTCCCTCACCTATCCCTAAACAGGCTGTAGATTAACAAACAGGATTAATTT 3060  
3118 CATCACTAAAGAAATCTTATTAATCACTGCTGCTTCCAGAGACTTTTCTCTAGAGCGG 3177  
3061 CATCACTAAAGAAATCTTATTAATCACTGCTGCTTCCAGAGACTTTTCTCTAGAGAGCG 3120

## RESULT 14

US-08-446-648-22

; Sequence 22, Application US/08446648

; Patent No. 6331302

; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.

; APPLICANT: Bennett, Brian D.

; APPLICANT: Goeddel, David

; APPLICANT: Lee, James M.

; APPLICANT: Matthews, William

; APPLICANT: Teai, Siao Ping







Search completed: August 28, 2003, 05:51:59  
Job time : 293.45 secs

QY	2038	TGCACACTGTGTCAGGACCAATTTTACTGTGATTTTGTGAAATCTCTGTGCTATGFGTATCTTCTC	2155
DB	2041	TGCACACTGTGTCAGGACCAATTTTACTGTGATTTTGTGAAATCTCTGTGCTATGFGTATCTTCTC	2100
QY	2158	AACATCTCTAAGAAGTAAAGAGAGAAAATTTCCACAGGACTTGGACAGAGAGATTTTTCACAGGAA	2217
DB	2101	AACATCTCTAAGAAGTAAAGAGAGAAAATTTCCACAGGACTTGGACAGAGATTTTTCACAGGAA	2160
QY	2218	CACAAATTTCAAGTTTATACCCCACTTTTCCAAATCACATCCAAATTTCCAGCATGCCTGGTTCA	2277
DB	2161	CACAAATTTCAAGTTTATACCCCACTTTTCCAAATCACATCCAAATTTCCAGCATGCCTGGTTCA	2220
QY	2278	AGAGAGATTTCAGATACACCCCGGACTCGGATCAAAATCTCAGGGCTTCATGCGAAATTCATTT	2337
DB	2221	AGAGAGATTTCAGATACACCCCGGACTCGGATCAAAATCTCAGGGCTTCATGCGAAATTCATTT	2280
QY	2338	CACCTCTGAAGATGAAATTTGAATATGAAACCACAAAAGGCTGGAGAGAGGAGGACTTTCG	2397
DB	2281	CACCTCTGAAGATGAAATTTGAATATGAAACCACAAAAGGCTGGAGAGAGGAGGACTTTCG	2340
QY	2398	AATGTGCTTACATTTTGAAGATCTTCTTTGCTTTCATATCAAGTTGTCACCAAGGAATGGAA	2457
DB	2341	AATGTGCTTACATTTTGAAGATCTTCTTTGCTTTCATATCAAGTTGTCACCAAGGAATGGAA	2400
QY	2458	TTTCTCGGAATTTAAGTCGTGTGTTCACAGAGACTGCGCCGCCAGGAACGTGCTGTGCACC	2517
DB	2401	TTTCTCGGAATTTAAGTCGTGTGTTCACAGAGACTGCGCCGCCAGGAACGTGCTGTGCACC	2460
QY	2518	CACGGGAAAGTGTGTAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTGATTC	2577
DB	2461	CACGGGAAAGTGTGTAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTGATTC	2520
QY	2578	AACATATGTTGTACGGGGCAATGCCCCTGTGCTGTGAAATTCGATGGGCCCCGAGAAAGCCCTG	2637
DB	2521	AACATATGTTGTACGGGGCAATGCCCCTGTGCTGTGAAATTCGATGGGCCCCGAGAAAGCCCTG	2580
QY	2638	TTTGAAGGCATCTACACCAATTAAGAGTGATGTCTGTGTCATATGGAAATTTACTGTGGGAA	2697
DB	2581	TTTGAAGGCATCTACACCAATTAAGAGTGATGTCTGTGTCATATGGAAATTTACTGTGGGAA	2640
QY	2698	ATCTTCTCATTGTGTGTGAATCTTTACCTTGGCAATTCCTGGTTGATGTGATGCTTCTACAAA	2757
DB	2641	ATCTTCTCATTGTGTGTGAATCTTTACCTTGGCAATTCCTGGTTGATGTGATGCTTCTACAAA	2700
QY	2758	CTGATTCAAAATGGAATTTAAAATGGATCGCCATCTTATGCTACAGAGAAATATACATTT	2817
DB	2701	CTGATTCAAAATGGAATTTAAAATGGATCGCCATCTTATGCTACAGAGAAATATACATTT	2760
QY	2818	ATAATGCAATCTCTCTGGGCTTTTGTACTAAGGAAACGGCCATCTTCTCCCTAATTTGACT	2877
DB	2761	ATAATGCAATCTCTCTGGGCTTTTGTACTAAGGAAACGGCCATCTTCTCCCTAATTTGACT	2820
QY	2878	TCGTTTTTAGATGTGAGCTGGCAGATGCGAAGAAAGCGATGATCAGAAATGTGGATGGC	2937
DB	2821	TCGTTTTTAGATGTGAGCTGGCAGATGCGAAGAAAGCGATGATCAGAAATGTGGATGGC	2880
QY	2938	CGTGTCTTGGAAATGTCCTCTCACCTACCCAAAACAGGGGACCTTTTCAGCAGAGATGGAT	2997
DB	2881	CGTGTCTTGGAAATGTCCTCTCACCTACCCAAAACAGGGGACCTTTTCAGCAGAGATGGAT	2940
QY	2998	TTGGGGCTACTCTCTCCGCAAGGCTCAGGTCGAAAGATTCGTAGAGGAAACAAATTTAGTTTTA	3057
DB	2941	TTGGGGCTACTCTCTCCGCAAGGCTCAGGTCGAAAGATTCGTAGAGGAAACAAATTTAGTTTTA	3000
QY	3058	AGGACTTATCCCTCCACTATCCCTAACAGGCTGTAGATTACCAAAAACAGATTAATTT	3117
DB	3001	AGGACTTATCCCTCCACTATCCCTAACAGGCTGTAGATTACCAAAAACAGGTTAAITTT	3060
QY	3118	CATCATTAAGAAATCTATATCACTGCTGCTTCCACGACTTTTCTCTAGAGAGCG	3177
DB	3061	CATCATTAAGAAATCTATATCACTGCTGCTTCCACGACTTTTCTCTAGAGAGCG	3120





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2003, 07:20:58 ; Search time 33.0166 Seconds  
(without alignments)  
2892.346 Million cell updates/sec

Title: US-09-919-408A-4

Perfect score: 5274

Sequence: 1 MPALARDAGTVPLLVVFSAM.....PFSREMDLGLSPQAQVEDS 993

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76:\*

1: Piri:\*

2: Pir2:\*

3: Pir3:\*

4: Pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	5238	99.3	993	2	A36873
2	4533.5	86.0	1000	2	S18827
3	4429.5	84.0	992	2	A39931
4	1286	24.4	977	2	I45877
5	1266	24.0	975	1	TVMSKT
6	1263.5	24.0	978	1	A49814
7	1251	23.7	976	1	TVHUKT
8	1226	23.2	954	2	I51703
9	1225.5	23.2	980	1	TVCTMD
10	1218	23.1	941	1	TVVMMD
11	1216.5	23.1	972	1	TVHUMD
12	1215	23.0	960	1	JN0877
13	1213	23.0	975	2	T30816
14	1194	22.6	978	2	S16385
15	1183.5	22.4	976	1	TVMSMD
16	1176	22.3	1088	1	PFRTGA
17	1166	22.1	1089	1	PFHUGA
18	1146	21.7	1089	1	S33727
19	1128	21.4	1087	2	I51552
20	1105.5	21.0	1098	1	PFMSRB
21	1086.5	20.6	1106	1	PFHUGB
22	1080	20.5	1048	2	T30815
23	992	18.8	790	1	F0W9H2
24	990	18.8	1338	2	S09982
25	959.5	18.2	1333	2	I78875
26	956	18.1	1336	2	I60598
27	944.5	17.9	1356	2	JC1402
28	939.5	17.8	1330	2	S49010
29	935	17.7	1348	2	S51656

30	932	17.7	1379	2	JC4954	vascular endotheli
31	923	17.5	1367	2	A41228	protein-tyrosine k
32	907.5	17.2	1363	2	I58375	protein-tyrosine k
33	904.5	17.2	1298	2	A48999	protein-tyrosine k
34	874.5	16.6	823	2	E35963	protein-tyrosine k
35	851	16.1	160	2	A39061	protein-tyrosine k
36	849	16.1	821	1	TVHUP2	fibroblast growth
37	842	16.0	813	1	A49123	fibroblast growth
38	841	15.9	824	2	S24108	protein-tyrosine k
39	836.5	15.9	822	2	A45081	fibroblast growth
40	832.5	15.8	824	2	S36439	fibroblast growth
41	828.5	15.7	769	2	S16236	fibroblast growth
42	828.5	15.7	822	2	A41794	keratinocyte growt
43	827	15.7	797	2	S38579	fibroblast growth
44	825	15.6	821	1	TVMSBK	fibroblast growth
45	823.5	15.6	822	2	B54846	fibroblast growth

ALIGNMENTS

RESULT 1

A36873

protein-tyrosine kinase (EC 2.7.1.112) STK-1 precursor - human

N;Alternate names: stem cell tyrosine kinase 1

C;Species: Homo sapiens (man)

C;Date: 03-Nov-1995 #sequence\_revision 03-Nov-1995 #text\_change 04-Feb-2000

C;Accession: A36873

R;Small, D.; Levenstein, M.; Kim, E.; Carow, C.; Amin, S.; Rockwell, P.; Witte, L.; Bur

Proc. Natl. Acad. Sci. U.S.A. 91, 459-463, 1994

A;Title: STK-1, the human homolog of Flk-2/Flt-3, is selectively expressed in CD34(+) h

A;Reference number: A36873; MUID:94119906; PMID:7507245

A;Accession: A36873

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-993 <MA>

A;Cross-references: GB:U02687

A;Note: in the authors translation, an additional residue Ala is shown after 420-Ala an

C;Genetics:

A;Map position: 13q12

C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol.

C;Keywords: ATP; glycoprotein; phosphotransferase; transmembrane protein; tyrosine-spec

F;608-950/Domain: protein kinase homology <Kin>

F;616-624/Region: protein kinase ATP-binding motif

Query Match 99.3%; Score 5238; DB 2; Length 993;

Best Local Similarity 99.7%; Pred No. 4.2e-248;

Matches 991; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 1 MPALARDAGTVPLLVVFSAMIFGTITNQDLFVVKVILNHKNNDSSVYKSSSYPMVSESP 60

DB 1 MPALARDAGTVPLLVVFSAMIFGTITNQDLFVVKVILNHKNNDSSVYKSSSYPMVSESP 60

QY 61 EDLGCALRQSSGTVYEAARVVDVVSASITQVLVDAPGNISCLWPKHSLNCQHPFDL 120

DB 61 EDLGCALRQSSGTVYEAARVVDVVSASITQVLVDAPGNISCLWPKHSLNCQHPFDL 120

QY 121 QNRGVVSWILKMTQAGEYLLFIQSEATNYILFTVSRNTLLYTLRPPYFRKMENQD 180

DB 121 QNRGVVSWILKMTQAGEYLLFIQSEATNYILFTVSRNTLLYTLRPPYFRKMENQD 180

QY 181 ALVCISESPEPIVWVLCDQSQESCKESPAVVKKEKVLHFLFGTDIRCCARNELGRE 240

DB 181 ALVCISESPEPIVWVLCDQSQESCKESPAVVKKEKVLHFLFGTDIRCCARNELGRE 240

QY 241 CTRLFTIDLNQTPQTLPLQFLKVGEPILWIRKCAVHNHGFGLTWELNKALEEGNYFEM 300

DB 241 CTRLFTIDLNQTPQTLPLQFLKVGEPILWIRKCAVHNHGFGLTWELNKALEEGNYFEM 300

QY 301 STYSTNRTMIRILFAFVSSVARNDTCYTCSSSKHPSQSALVTIVGKGPINATNSSDYDE 360

DB 301 STYSTNRTMIRILFAFVSSVARNDTCYTCSSSKHPSQSALVTIVGKGPINATNSSDYDE 360

QY	361	IDQYEFPCSVRKPAYPOIRCTWTFRRKSPCKGKLDNGYSISKCNHKGHPGEYIFH--	419
Db	361	IDQYEFPCSVRKPAYPOIRCTWTFRRKSPCKGKLDNGYSISKCNHKGHPGEYIFHA	420
QY	420	AENDDAQFTKMTLNTRRKPOVLAEASASQSCFSDGYPLPSMTWKCKSDKSPNCTEEIT	479
Db	421	AENDDAQFTKMTLNTRRKPOVLAEASASQSCFSDGYPL-SMTWKCKSDKSPNCTEEIT	479
QY	480	EGVWNRKANRKVFGQWSSSTLNMSBAIKGFLVKCCAYNSLGTSCETILNLSGPPFPFIQ	539
Db	480	EGVWNRKANRKVFGQWSSSTLNMSBAIKGFLVKCCAYNSLGTSCETILNLSGPPFPFIQ	539
QY	540	DNISFYATTGCVLLFIWLTLLICHKYKQPRYESQLOMVOVTGSSDNEYFYVDPRYEFY	599
Db	540	DNISFYATTGCVLLFIWLTLLICHKYKQPRYESQLOMVOVTGSSDNEYFYVDPRYEFY	599
QY	600	DLKWEFPRENLEFGKVLGSGAFGKVMNATAYGSKTGVSIOVAVKMLKEKADSSEREA	659
Db	600	DLKWEFPRENLEFGKVLGSGAFGKVMNATAYGSKTGVSIOVAVKMLKEKADSSEREA	659
QY	660	SELKMMTOLGSHENINVLGACTLSGPIYLIFBYCCYGDLLNLYLSRKREKPHRTWTEIFK	719
Db	660	SELKMMTOLGSHENINVLGACTLSGPIYLIFBYCCYGDLLNLYLSRKREKPHRTWTEIFK	719
QY	720	EHNFSYPTQSHPNSSMPGSRVQIHPDSDQISGLHNSFHSEDEIEYENOKRLEEEED	779
Db	720	EHNFSYPTQSHPNSSMPGSRVQIHPDSDQISGLHNSFHSEDEIEYENOKRLEEEED	779
QY	780	LNVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIMSD	839
Db	780	LNVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIMSD	839
QY	840	SNYVVRGNARLPVKWMAPESLFEGITYIKSDVMSYGILLWEIFSLGVNYPGIPVDANFY	899
Db	840	SNYVVRGNARLPVKWMAPESLFEGITYIKSDVMSYGILLWEIFSLGVNYPGIPVDANFY	899
QY	900	KLIONGFMDQOPFYATEEIIYIMOSWAFDSRKRPSPNLTSLFLGCOLADAEAMYNQVD	959
Db	900	KLIONGFMDQOPFYATEEIIYIMOSWAFDSRKRPSPNLTSLFLGCOLADAEAMYNQVD	959
QY	960	GRVSECPHTYONRRPSPREMDLGLLSPQAQVEDS	993
Db	960	GRVSECPHTYONRRPSPREMDLGLLSPQAQVEDS	993
RESULT 2			
S18827			
F1c3 protein - mouse			
C:Species: Mus musculus (house mouse)			
C:Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 24-Sep-1999			
C:Accession: S18827			
R:Rosnet, O.; Marchetto, S.; deLapeyriere, O.; Birnbaum, D.			
Oncogene 6, 1641-1650, 1991			
A:Title: Murine F1c3, a gene encoding a novel tyrosine kinase receptor of the PDGFR/CSF1			
A:Reference number: S18827; MUID:92019834; PMID:1656368			
A:Accession: S18827			
A:Molecule type: mRNA			
A:Residues: 1-1000 <ROS>			
A:Cross-references: EMBL:X59398; NID:g50978; PIDN:CAA42041.1; PID:g50979			
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo			
F:609-953/Domain: protein kinase homology <KIN>			
F:617-625/Region: protein kinase ATP-binding motif			
Query Match 86.08; Score 4533.5; DB 2; Length 1000;			
Best Local Similarity 85.66; Pred. No. 9.3e-214;			
Matches 851; Conservative 56; Mismatches 84; Indels 3; Gaps 2;			
QY	1	MPALA-RDAGTVPLLVFSAIMFCTITNQDLPVIKCVLIHNKNDSSGVKSSSSYPMVSES	59
Db	1	MRALAQRSDRLLLVVLVSLVILETVITNQDLPVIKCVLIHNKNDSSGVKSSSSYPMVSES	60
QY	60	PEDLGCAIRPQSSGTGYTAAAEVDSASITLQVLVDAPGNISCLWVFKHSLNLCOPHFD	119

Db	61	PEDLQCTPRQSEGTYEATVEAESGSITLQVQLATFGDLSCLMVFKHSSLCQPHFD	120
QY	120	LQNRGVSVVILKMTETOAGEYLLFTQSEATNTIILFTVISIRNTLYTLRRPFRKXENQ	179
Db	121	LQNRGVSMAILNVETQAGEYLLHTQSEAAANTVLTFTVNRDTQTLVLRPFRKXENQ	180
QY	180	DALVCISESVPPIVEMVLCDSQESCKEESPAVVKKEKVLHELFGDTRCCARNELGR	239
Db	181	DALLCISEGVPEPTEVWILCSSHRESCKEESPAVVRKEKVLHELFGDTRCCARNALGR	240
QY	240	ECTRLFTIDLNPOTTLTLPQLFKVGEPLWIRKCAVHVNHGFGILTWELNKALEEGNYFE	299
Db	241	ESTKLEFTIDLNQAQSTLTPQLFKVGEPLWIRKCAVHVNHGFGILTWELNKALEEGSYFE	300
QY	300	MSTYSTNRTMIRILPAFVSVSARNDGYTCTSSSKHPSQSALVTIIVKGFINATNNSSEY	359
Db	301	MSTYSTNRTMIRILLAFVSVSGRNDGYTCTSSSKHPSQSALVTILEKGINATSSOEY	360
QY	360	EIDQYEFPCSVRKPAYPOIRCTWTFRRKSPCKGKLDNGYSISKCNHKGHPGEYIFH	419
Db	361	EIDPYEKFCFSVRKPAYPRIRCTWIFSOAFPCQKRGLEDGYISIKPCDHKNKPGYIFY	420
QY	420	AENDDAQFTKMTLNTRRKPOVLAEASASQSCFSDGYPLPSMTWKCKSDKSPNCTEEIT	479
Db	421	AENDDAQFTKMTLNTRRKPOVLAEASASQSCFSDGYPLPSMTWKCKSDKSPNCTEEIP	480
QY	480	EGVWNRKANRKVFGQWSSSTLNMSBAIKGFLVKCCAYNSLGTSCETILNLSGPPFPFIQ	539
Db	481	EGVWNRKANRKVFGQWSSSTLNMSBAIKGFLVKCCAYNSMTGTCETIFLNSGPPFPFIQ	540
QY	540	DNISFYATTGCVLLFIWLTLLICHKYKQPRYESQLOMVOVTGSSDNEYFYVDPRYEFY	599
Db	541	DNISFYATTGCVLLFIWLTLLICHKYKQPRYESQLOMVOVTGSSDNEYFYVDPRYEFY	600
QY	600	DLKWEFPRENLEFGKVLGSGAFGKVMNATAYGSKTGVSIOVAVKMLKEKADSSEREA	659
Db	601	DLKWEFPRENLEFGKVLGSGAFGKVMNATAYGSKTGVSIOVAVKMLKEKADSSEREA	660
QY	660	SELKMMTOLGSHENINVLGACTLSGPIYLIFBYCCYGDLLNLYLSRKREKPHRTWTEIFK	719
Db	661	SELKMMTHLGHDNINVLGACTLSGPIYLIFBYCCYGDLLNLYLSRKREKPHRTWTEIFK	720
QY	720	EHNFSYPTQSHPNSSMPGSRVQIHPDSDQISGLHNSFHSEDEIEYENOKRLEEEED	777
Db	721	EHNFSYPTQSHPNSSMPGSRVQIHPDSDQISGLHNSFHSEDEIEYENOKRLEEEED	780
QY	778	EDLNVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIM	837
Db	781	EDLNVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL	840
QY	838	SDSNVYVRGNARLPVKWMAPESLFEGITYIKSDVMSYGILLWEIFSLGVNYPGIPVDAN	897
Db	841	SDSNVYVRGNARLPVKWMAPESLFEGITYIKSDVMSYGILLWEIFSLGVNYPGIPVDAN	900
QY	898	FYKLIQNGFKMDQOPFYATEEIIYIMOSWAFDSRKRPSPNLTSLFLGCOLADAEAMYNQ	957
Db	901	FYKLIQSGFKMEQFPFYATEEIIYIMOSWAFDSRKRPSPNLTSLFLGCOLADAEAMYNQ	960
QY	958	VGRVSECPHTYONRRPSPREMDLGLLSPQAQVE	991
Db	961	MGNVPEHPSIYQNRRLPSREAGSEPPSPQAQVK	994
RESULT 3			
A39931			
protein-tyrosine kinase (EC 2.7.1.112) flk-2, hematopoietic - mouse			
C:Species: Mus musculus (house mouse)			
C:Date: 17-Jan-1992 #sequence_revision 17-Jan-1992 #text_change 04-Feb-2000			
C:Accession: A39931			
R:Matthews, W.; Jordan, C.T.; Wiegand, G.W.; Pardoll, D.; Lemischka, I.R.			
Cell 65, 1143-1152, 1991			
A:Title: A receptor tyrosine kinase specific to hematopoietic stem and progenitor cell-			

A:Reference number: A39931; MUID:91292518; PMID:1648448

A:Accession: A39931

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-932 <MAT>

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homology

C:Keywords: ATP; phosphotransferase; transmembrane protein; tyrosine-specific protein kinase

F:609-953/Domain: protein kinase homology <KIN>

F:617-625/Region: protein kinase ATP-binding motif

Query Match 84.0%; Score 4429.5; DB 2; Length 992;

Best Local Similarity 84.1%; Pred. No. 1.1e-208;

Matches 836; Conservative 57; Mismatches 90; Indels 11; Gaps 4;

QY 1 MPALA-RDAGTVPLLVFAMIGTITNQDLPVVKVILNHKNQSSVQKSSSYPMVSS 59  
 DB 1 MRALQSRDRLLLVLSMILETVNQDLPVVKVILSHENNGSSAGKPSYMRVGS 60  
 QY 60 PEDLGCALRPOSSGTVYEAADVVSASITLQVLVDAPGNISCLWVFKHSLNCPHED 119  
 DB 61 PEDLQCTPRQSEGTVYEAADVVSASITLQVLVDAPGNISCLWVFKHSLNCPHED 120  
 QY 120 LQNRGVSMVLKMTQAGEYLLFTQSEATNYTLFTVSIIRNTLYTLRRPVFRKMNQ 179  
 DB 121 LQNRGIVSMALLNVETQAGEYLLHIOSEBANYTLFTVNVHDTQLVLRPVFRKMNQ 180  
 QY 180 DALVCISEVPPIVWVLCDQSGESCKEESPAVVKKEKVLHFLPGTDIRCCARNELGR 239  
 DB 181 DALLCISEGPEPTVWVLCSHRESCKEESPAVVKKEKVLHFLPGTDIRCCARNALGR 240  
 QY 240 ECTRFLTIDNQPTTLTLPOLFLKVGEPVIRKCAVHNVHGLTWELNKALEEGNYFE 299  
 DB 241 ECKTLFTIDLQAPOSTLPOLFLKVGEPVIRKCAVHNVHGLTWELNKALEEGNYFE 300  
 QY 300 MTSYSTNRTMIRILFAFVSVAARNDTGYTCSSKHPQSALVTIVKGFINATNSDEY 359  
 DB 301 MTSYSTNRTMIRILFAFVSVAARNDTGYTCSSKHPQSALVTILEKGFINATNSDEY 360  
 QY 360 EIDYVEFCFSVRKAYPOIRCTWTFSRKSPCEQKGLDNGYSISKFCNHKHPQGEYI 419  
 DB 361 EIDYVEFCFSVRKAYPRICTWTFSPQASFPCEQKGLDNGYSISKFCNHKHPGEYI 420  
 QY 420 AENDDAQFTMTLIRKPOVLAESAQSCFSDGYPLPWTWKCKSDKSPNCTEET 479  
 DB 421 AENDDAQFTMTLIRKPOVLAESAQSCFSDGYPLPWTWKCKSDKSPNCTEET 480  
 QY 480 EGVWNRKANRVKFGWVSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPPFIQ 539  
 DB 481 EGVWNRKANRVKFGWVSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPPFIQ 540  
 QY 540 DNISFYATIGVCLLFTVLTLAI CHYKKQFRYESQLOMVOVQTGSSDNDEYFYVDREY 599  
 DB 541 DNISFYATIGVCLLFTVLTLAI CHYKKQFRYESQLOMVOVQTGPDLDNEFYVDREY 600  
 QY 600 DLKWFPPRENLEFGKVLGSGAFKGMNATAYGISTGVSIVQAVXMKLKEKADSSREALM 659  
 DB 601 DLKWFPPRENLEFGKVLGSGAFKGMNATAYGISTGVSIVQAVXMKLKEKADSSREALM 660  
 QY 660 SELKMTQLGSHENIVNLLGACTLSGPVILIPEYCCYGLLNLVLRSKREKPHRTWTEIFK 719  
 DB 661 SELKMTQLGSHENIVNLLGACTLSGPVILIPEYCCYGLLNLVLRSKREKPHRTWTEIFK 720  
 QY 720 ENHFSFYPTFQSHNSMPGSREVQIHPDSDOI SGLHGNFSHSEDBIEYENOKRL--EE 777  
 DB 721 ENHFSFYPTFQSHNSMPGSREVQIHPDSDOI SGLHGNFSHSEDBIEYENOKRLAEE 780  
 QY 778 EDLNVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHKVKVICDFGLARDIM 837  
 DB 781 EDLNVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHKVKVICDFGLARDIL 840  
 QY 838 SDSNYVVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIFSLGYNPYPGIPVDAN 897

DB 841 SDSYVVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIFSLGYNPYPGIPVDAN 900  
 QY 898 FYKLQNGFKMDOPFYATEEIIYINOSWAPDSRKRPSFPNLTSLGCOLAEEAMYQN 957  
 DB 901 FYKLQNGFKMDOPFYATEEIIYINOSWAPDSRKRPSFPNLTSLGCOLAEEAMYQN 957  
 QY 958 VDRVSECPHTYQNRPPFSREMDLGLLSPQAQVE 991  
 DB 958 ----IRTSIHLPKQAAPQORG-GLRAQSPQOVK 986

RESULT 4

I45877

protein-tyrosine kinase (EC 2.7.1.112), receptor type kit precursor - aurochs

C:Species: Bos primigenius (aurochs)

C:Date: 19-Dec-1997 #sequence revision 19-Dec-1997 #text\_change 20-Apr-2000

C:Accession: I45877

R:Kubota, T.; Hikono, H.; Sasaki, E.; Sakurai, M.

Gene 141, 305-306, 1994

A:Title: Sequence of a bovine c-kit proto-oncogene cDNA.

A:Reference number: I45877; MUID:94215924; PMID:7512939

A:Accession: I45877

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-977 <KUB>

A:Cross-references: GB:D16680; MID:9516659; PIDN:BAA04084.1; PID:9516660

C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology.

C:Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase

F:329-395/Domain: immunoglobulin homology <IMM>

F:588-932/Domain: protein kinase homology <KIN>

Query Match 24.4%; Score 1286; DB 2; Length 977;

Best Local Similarity 33.2%; Pred. No. 1.8e-55;

Matches 332; Conservative 172; Mismatches 344; Indels 152; Gaps 36;

QY 47 VCKSSYPWVSSPDLGC-ALRPOSSGTVYEAADVVSASITLQVLVDAPGNISCLW 105  
 DB 20 VOTGSSQPSV--SPGELSPLSHPAKS-----ELIVSGDEIRLCTDPGFVK--W 66  
 QY 106 VPKHSSLNCPHFIDLQNRGVSMVLKMTQAGEYLLFIQSEAT---NYTLFTVSI 162  
 DB 67 IFE-----ILQSLSEKTNPEMIT-EKAEATNTGTYCTNKGGLSS 105  
 QY 163 TL-----LYTURLPYFRMKNQDALVCISEVPPIVWVLCDQSGESCKEESPAV 213  
 DB 106 SIYFVVRDPEKFLIDLPLYGKEEN-DTLVRCPLTDPE-VTNYSLTGCEGKPLKPLT 163  
 QY 214 -----VKKEEKVHLHFLGTDIRC--CARNELGRE-CTRLFTID----LNQTPQT 258  
 DB 164 ADPKAGITIRNVKREYHRL-----CLHCSANQKSKMLSKKFTLLKVRRAAIKAVPV 217  
 QY 259 Q--LFLKVGEPVIRKCAVHNVHGLTWELNKALEEGNYPEMSTYSTNRTMIRILFA 316  
 DB 218 KTSYLLRGEPEFAVTCIKDVSSVSDSMWIKENSQOTKATKNSHQDFFYLROERLT 277  
 QY 317 VSSVARNDTGYTCSSKHPQSALVT---IVKGFINA-----TNSEDEYIDQY 364  
 DB 278 ISSARVNDSGVPMCVANNTFGSANVTTLVVVDKGFINIFPMNMTTVFVNDGENVDL 334  
 QY 365 BEFCFSVFKAYPO-IRCTWTFSRKS-----PFCEQKGLDNGYSISKFCNHK 411  
 DB 335 -----VVEYEAPKPVHRQWIYMNRTSTDKWDDYPKSE-----NESNIRYVNLHL 384  
 QY 412 --OPGEYILFAENDDAQFTKMTLIRKPOVLAESA--SQASCFSDGYPLPWSMTWK 467  
 DB 385 GTEGTYTFHVSNSDVNSVTFNVNTKPELLTHDLVNGMLQCVAAGFPETDIDWYFC 444  
 QY 468 SDKSPNCTEEI--TEGVWNRKANRVKFGWVSSSTLNMSAIGFLVKCCAYNSLGTSC 526  
 DB 445 PTEQRCSVPVGPVDVQIQNSSVSPFKLVVYSTIDSTFFKNGTVECRAYNDVGKSAS 504  
 QY 527 ILLNSPGPPP-----FIQDNISFYATIGVCLLFTVLTLAI CHYKKQFRYESQLOM 578



788 LLCFAVOAKGMEFLBPKSVHRDLAARNVLVTHGKWKICDPGLARDIMSDSNVYVRGN 847  
 767 LLSFSQVAKAKAFASKNCIHRDLAARNLLTHGRITKICDPGLARDINDSNVYVRGN 826  
 848 ARLPVKWMAPELFGIYTIKSDVWSYGILLWEIFSLGVNYPGIPVDANFYKLIQNGFK 907  
 827 ARLPVKWMAPELFGIYTIKSDVWSYGIFLWELSLGSSPYGMPVDSKFKYKMIKEGFR 886  
 908 MDQPYATEEIIIMOSCHAFDSRKPSPNLTSLFGCCQLADAEAEAMQNV 958  
 887 MVSPEHAPAEYDMVKTCWDADPLKEPTKQVQVQLIEKQISDSTKHIYNSL 937

RESULT 6  
 A49814  
 protein-tyrosine kinase (EC 2.7.1.112), receptor type kit precursor - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Jun-2000  
 C:Accession: A49814; S49088  
 R:Tsujiyama, T.; Hirota, S.; Nomura, S.; Niwa, Y.; Yamazaki, M.; Tono, T.; Morii, E.; Kikuchi, T.; Tsujiyama, T.; Tono, T.; Yamazaki, M.; Nomura, S.; Kitamura, Y.  
 submitted to the EMBL Data Library, October 1991  
 A:Description: Two isoforms of rat c-kit receptor tyrosine kinase.  
 A:Reference number: S49088  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-511, 516-978 <TSU>  
 A:Cross-references: EMBL:62491; NID:G509135; PIDN:CAA4354.1; PID:G509136  
 C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;  
 C:Keywords: ATP; autophosphorylation; phosphotransferase; tyrosine-specific protein kinase  
 F:129-188/Domain: immunoglobulin homology <IMM>  
 F:589-932/Domain: protein kinase homology <KIM>  
 F:597-605/Region: protein kinase ATP-binding motif

Query Match 24.0%; Score 1263.5; DB 1; Length 978;  
 Best Local Similarity 33.8%; Pred. No. 2.3e-54;  
 Matches 334; Conservative 171; Mismatches 361; Indels 121; Gaps 37;

35 CVLHNKNDSSVGKSSSYPMVSESPEDLGCALRPOSSGTVYEAADVDSASITLQVL 94  
 12 CVLLVLLRQGTGTQSPSAGSPSP-----SIQPAQS-----ELIVEAGDTIRLT 57  
 95 VDAPGNISCLWVFKHSLNQCQPHDLQNRGVGSMVILKQTE-TOAGEYLLFIQSEATNYT 153  
 58 CTDPAFVK--WTFE-----ILDVRIENKQSEWREKAEATHQKYTCVSSGLRSSI 107  
 154 ILFTVSIIRN-TLTYLIRRYFRKMNQDALVCISVPPIVFWLVCDSQGSCKEESPA 212  
 108 YVFE---VRDPAVFLVGLPLFGK-EDNDALVRCLPTDPO-VSNYSLIECDGKSLPTDLKF 162  
 213 V-----VKEEKVLHFLGTDIRCCARNELGRECTRLFTID----LNQTPQTLPLQ 260  
 163 VFNPKAGITIKNVKRAYHRLC---IRCAQREGKWMRSDFTLKVRAAIKAFVWSVPET 219  
 261 --FLKVGPELWIRCKAVHVNHGFLTW-----ELENALEEGNVFEMS--TYSTNRMTIR 311  
 220 SHLLKEGDTFTVCTIKDVTSTVDSVMWIKLNPQSKAQVKNRWSWQGFNRYERQETLT- 278  
 312 ILFAFVSSVARNDTGYTTCSSSRHPQSALVT---IVGKGFN---ATNSSEYIDQYE 365  
 279 -----ISSARNVDSGVFMCYANNFTGSANVTTLKVEKGFNIFPVKNIT--VFVTDGE 331  
 366 EFCFSVRKAYPO-IRCTWTFSRKSFPCEQKGLD-----NGYSISKPCNHHK-----QP 413

332 NVDLVVEFEAYPKPEHQWIYMNRT--PTNRGEDYVKSNDQSNIRYVNELELRLTKGTEG 389  
 414 GEYIFHAENDDAQFTQWFTLNIRPKPOVLA--EASASQASCFSDGYPLPGLPWTWKCKSDKS 471  
 390 GYTFVLVNSDVSASVTFDVYVNTKPEILTYDRLMNGLRQCVAAAGPPEPTIDYDFCTGAE 449  
 472 PNCTEITE-GVNNRKANRKFQGVWSSSTLNNSEAIKGLFKVCCAYNSLSTCETILLN 530  
 450 QRCITVPVPDVQIQONASVSPFGKLVVQSSIDSSVFRHNGTVECKASNAVGS--SAFFN 507  
 531 SPGPPPF-----IDN-----ISFYATIGVCLLFIIVLTLLCHTKYKQFRIESQL 576  
 508 ----FAFKGNSKEIQIOPHTLFTPLIGFVVVTA--LMGIIVMVL--AYKYLQPMYEVQW 559  
 577 OMV-QVTGSSDNEYFYVDFREYEDLKWEFPRENLEFGVLGSGAGKWNATAYGISKT 635  
 560 KVVEING---NNYVYIDPTQLPYDHKWEFPNRNLSFGKTLGAGAFGVVEATAYGLIKS 616  
 636 GVSIOVAVKMLKEKADSSREALMSELKQMTQLGSHENIVNLIGACTLSPGIYLIPEYCC 695  
 617 DAAMTVAVKMLPESAHLTREALMSELKVLVSLGNHNVNLLGACTVGGPTLVITEYCC 676  
 696 YGDLNLYLSKREKFRHTWTEIFKEHNFYPTFQSHPNSSMPGSRB-VQIHPSDDOISG 754  
 677 YGDLNLYLRKQDSF--IPSKOBEQADALYKXNLLHSKSSCDSSNEYMDKPGVSYVVP 734  
 755 LHGNSPHS---EDEIYENOKLEEEEDLNLTFLFCFAYOVAKGMEFLPKSCVHRD 811  
 735 TKTDKERSARISYIERDVTYPAIMEDDEL-ALDLELLSFSYQVAKGMFLASKNCIHRD 793  
 812 LAARNVLVTHGKWKICDPGLARDIMSDSNVYVRGNARLPVKWMAPELFGIYTIKSDV 871  
 794 LAARNILLTHGRITKICDPGLARDINDSNVYVRGNARLPVKWMAPELFGIYTIKSDV 853  
 872 WSYGILLWEIFSLGVNYPGIPVDANFYKLIQNGFKMDQPFYATEEIIIMOSCHAFDSR 931  
 854 WSYGIFLWELFSLGSSPYGMPVDSKFKYKMIKEGFRMLSPHAPAEYEVKTCWDADPL 913  
 932 KPSFNLTSFLGCLADAEAEAMQNV 958  
 914 KRPTKQVQVQLIEKQISDSTKHIYNSL 940

RESULT 7  
 TVHUKT  
 protein-tyrosine kinase (EC 2.7.1.112), receptor type kit precursor - human  
 N:Alternate names: mast/stem cell growth factor receptor; tyrosine kinase receptor c-kit  
 C:Species: Homo sapiens (man)  
 C:Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 21-Jul-2000  
 C:Accession: S01426; PC1015; A41815; B41815; I37948; I56954; I54336  
 R:Yarden, Y.; Kuang, W.J.; Yang-Feng, T.; Coussens, L.; Munemitsu, S.; Dull, T.J.; Chen  
 EMO J. 6, 3341-3351, 1987  
 A:Title: Human proto-oncogene c-kit: a new cell surface receptor tyrosine kinase for an  
 A:Reference number: S01426; MUID:88111521; PMID:2448137  
 A:Accession: S01426  
 A:Molecule type: mRNA  
 A:Residues: 1-976 <YAR>  
 A:Cross-references: GB:X06182; NID:G34084; PIDN:CAA29548.1; PID:G34085  
 R:Hu, W.X.; Cornu, F.; Andre, C.; Galibert, F.  
 Chinese Biochem. J. 7, 618-629, 1991  
 A:Title: Nucleotide sequence of two neighbouring fragments of human c-kit proto-oncogene  
 A:Reference number: PC1015  
 A:Accession: PC1015  
 A:Molecule type: DNA  
 A:Residues: 412-713 <HUM>  
 A:Note: article in Chinese with English abstract  
 R:Spritz, R.A.; Giebel, L.B.; Holmes, S.A.  
 Am. J. Hum. Genet. 50, 261-269, 1992  
 A:Title: Dominant negative and loss of function mutations of the c-kit (mast/stem cell ;  
 A:Reference number: A41815; MUID:92133600; PMID:1370874  
 A:Accession: A41815  
 A:Molecule type: DNA  
 A:Residues: 579-583, 'L', 585-589 <SPR>

A;Cross-references: GB:S78839; NID:9244084; PIDN:AB21234.1; PID:9244085  
A;Note: sequence extracted from NCBI backbone (NCBIN:78839; NCBI:P:78842)  
A;Note: disease-related mutant from patient with piebaldism  
A;Accession: B41815  
A;Molecule type: DNA  
A;Residues: 637-641, 'SPELPW' <SP2>  
A;Cross-references: GB:S78843; NID:9244086; PIDN:AB21235.1; PID:9244087  
A;Note: sequence extracted from NCBI backbone (NCBIN:78843; NCBI:P:78844)  
A;Note: disease-related mutant from patient with piebaldism  
A;Accession: C41815  
A;Molecule type: DNA  
A;Residues: 556-560, 'GDKWK' <SP3>  
A;Cross-references: GB:S78845; NID:9244088; PIDN:AB21236.1; PID:9244089  
A;Note: sequence extracted from NCBI backbone (NCBIN:78845; NCBI:P:78846)  
A;Note: disease-related mutant from patient with piebaldism  
R;Gibbel, L.B.; Strunk, K.M.; Holmes, S.A.; Spritz, R.A.  
Oncogene 7, 2207-2217, 1992  
A;Title: Organization and nucleotide sequence of the human KIT (mast/stem cell growth factor receptor) gene  
A;Reference number: I37948; MUID:93064697; PMID:1279499  
A;Accession: I37948  
A;Molecule type: DNA  
A;Residues: 1-976 <RES>  
A;Cross-references: EMBL:X69301; NID:934089; PIDN:CAA9159.1; PID:9825686  
A;Note: an alternative splice form omitting residues 510-513 is described  
R;Yamamoto, K.; Tojo, A.; Aoki, N.; Shibuya, M.  
Jpn. J. Cancer Res. 84, 1136-1144, 1993  
A;Title: Characterization of the promoter region of the human c-kit proto-oncogene.  
A;Reference number: I56954; MUID:94103107; PMID:7506248  
A;Accession: I56954  
A;Molecule type: DNA  
A;Status: translated from GB/EMBL/DBJ  
A;Residues: 1-22 <RES>  
A;Cross-references: GB:S67773; NID:9459358; PIDN:AAB29529.1; PID:9459359  
R;Spritz, R.A.; Holmes, S.A.; Berg, S.Z.; Nordlund, J.J.; Fukai, K.  
Hum. Mol. Genet. 2, 1499-1500, 1993  
A;Title: A recurrent deletion in the KIT (mast/stem cell growth factor receptor) proto-oncogene  
A;Reference number: I54336; MUID:94061059; PMID:7694728  
A;Accession: I54336  
A;Molecule type: DNA  
A;Status: translated from GB/EMBL/DBJ  
A;Residues: 242-250 <RES>  
A;Cross-references: GB:S67686; NID:9460545; PIDN:AAD13996.1; PID:94261696  
C;Genetics:  
A;Gene: GDB:KIT  
A;Cross-references: GDB:120117; OMIM:164920  
A;Map position: 4q12-q12  
A;Introns: 23/1; 113/1; 207/1; 252/3; 309/1; 372/2; 411/1; 449/2; 514/1; 549/3; 592/1; 601/1  
A;Note: defects in this gene may result in piebaldism  
C;Function:  
A;Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP  
C;Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;  
C;Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; kinase-related  
protein; tyrosine-specific protein kinase  
F;1-976/Product: protein-tyrosine kinase kit precursor, long form #status predicted <MAT>  
F;1-509,514-976/Product: protein-tyrosine kinase kit precursor, short form #status predicted <SIG>  
F;1-22/Domain: signal sequence #status predicted <SIG>  
F;23-976/Product: protein-tyrosine kinase kit #status predicted <MAT>  
F;23-520/Domain: extracellular #status predicted <EXT>  
F;51-99/Domain: immunoglobulin homology <IMM1>  
F;129-188/Domain: immunoglobulin homology <IMM2>  
F;226-292/Domain: immunoglobulin homology <IMM3>  
F;328-394/Domain: immunoglobulin homology <IMM4>  
F;423-493/Domain: immunoglobulin homology <IMM5>  
F;521-543/Domain: transmembrane #status predicted <TM>  
F;544-976/Domain: intracellular #status predicted <INT>  
F;587-931/Domain: protein kinase homology <KIN>  
F;595-603/Region: protein kinase ATP-binding motif  
F;58-97, 136-186, 233-290, 428-491/Disulfide bonds: #status predicted  
F;130,145,283,293,300,320,352,367,463,486/Binding site: carbohydrate (Asn) (covalent) #  
F;623,640,792/Active site: Lys, Glu, Asp #status predicted  
F;797,810/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 23.7%; Score 1251; DB 1; Length 976;  
Best Local Similarity 32.7%; Pred. No. 9.2e-54;  
Matches 324; Conservative 166; Mismatches 352; Indels 148; Gaps 34;  
QY 47 VGKSSVPMVSESPEDLGCALRPOSSGTVEAAAVEVDVSASITLOVLVDAPGNISCLWV 106  
DB 20 VQTSSQPSVSPG-EPSPPSIHFGKSDLI-----VVGDEIRLLCTDTPGVK-WT 67  
QY 107 FKHSLLNCPHFDLQNRGVVSMVLKMTETQAGEYLLFIQSEATN---YTLFTVSIKNT 163  
DB 68 FE-----ILDETNEKNQEWIT-EKAEATNTKTYCTNKHGLSNS 106  
QY 164 L-----LYTLRRPYFRKMNQDALVCISSEVPPIVWVLCDSQGSCKEE----- 209  
DB 107 IYVVRDPAPKFLVDRLSYGK-EDNDTLVRCPLTDP-VTNYSLKGGKRLPKDLRFIP 164  
QY 210 ---SPAVVKKKEKVLHETFGDTRCCARNELGRECTLETFIDL---NQTPQTLPO--L 260  
DB 165 DPKAGIMIKSVKRAYHRLC---LHCSVDQEGKSVLSEKFLKVRPAKAVPVVSVSKASY 221  
QY 261 FLKVGEPILWIRCKAVHVNHGFLTWELN---KALBEGNYFEMSTYSTNRTMIRILFAFV 317  
DB 222 LLREGHEFTVCTIKDVSSSVYSTWKEENSQTKLQEKYNSMHHGDFNYERQAT-----LTI 277  
QY 318 SSVARNDTGYTCSSSKHPSQSALVT---IVGKGFNA-----TNSSEDEYIDQYE 365  
DB 278 SSARVNDSGVFMVCYANNTFGSANVTTLVVDVKGFINIPMINTVTFVNDGENDVL---- 333  
QY 366 EFCFSYRFKAYPO-IRCTWTFSRKSP--PCEQKGLDNGYSISKFCNKH-----QPGE 415  
DB 334 ----IVEYEAFFPKPEHQOQIYMNRTFTDKWEDYPKSESNIRYVSELHILTRLKGTGGT 389  
QY 416 YIFHAENDDAQTKMFTLNIRRPQVLA--EASASQASCFSDGVPPLPSWTWKKCSKSPN 473  
DB 390 YTLVNSDNDVNAIAFNYYVNTPEILTYDLVNGMLQCVAAAGEPEFIDWYFCPGTEQR 449  
QY 474 CTSEITE-GVMNRKANRKFQGVSSSTLAMSBAIKGLFKCCAYNSLGTSCETILNSP 532  
DB 450 CSASVLPVDVQTLNSSGPPFKLVQVSSIDSSAFKNGTVEKAYNDVGKT--SAYFN-- 505  
QY 533 GPFPFTQDN-----ISFYATIGVCLLFIWLTLLIICHKVKQKFRVESQLQM 578  
DB 506 --FAPKNNKEQIHPTLFTPLLGIFVIVAGMCIIVMLT---YKLOKPMYEVQMKV 559  
QY 579 V-QVTGSSDNEYFYVDPFREYEDLKWEFFRENLEFPGVLGSGAPGKVMNATAYGISKTV 637  
DB 560 VESING---NNYVIIDTQLPYDHKWEFPNRLSPGKTLGAGAFGVKVEATAYGLIKSDA 616  
QY 638 SIQVAVQMLKEKADSSEREALMSLKMNTOLGSHENIVNLLGACTLSGPIYLIFEYCCYG 697  
DB 617 AMTVAVQMLKPSAHLTEREALMSLKVLSYLGHNHNVNLLGACTIGGPTLVITEYCCYG 676  
QY 698 DILNLYRSKREKFRHWTE-----IPKEHNFYPTFQSHPNSSM---PGSREVOIHPD 748  
DB 677 DLLNFLRRKEDSFICSKQEDHAEALYKNLLHSEKSSCDSTNEYMDMKFGVSIV-VPTK 735  
QY 749 SDOISSLGHGSHSEDEIEYENKRLSEEDLNVLTPELILCFAYQVAKGMERLEPKSCV 808  
DB 736 ADKRSVRIGSY-----IERDVTAPMEDDEL-ALDLELLSFYQVAKGAFVLSKNCI 789  
QY 809 HRDLAARNVLVTHGKVKIKICDFGLARDIMSDSNVYVRGNARLPVKWMAPESLFEGITYIK 868  
DB 790 HRDLAARNILLTHGRITKICDFGLARDIKDSDNVYVKGARLPVKWMAPESFVNCVTFE 849  
QY 869 SDVWSYGILLWEIIFSLGVNPPYGPVDANFYKLIQNGFKWDQPPYATEEYIYINQSCWAF 928  
DB 850 SDVWSYGIFLWELFSLGSSPGMPGVDKSYKMKKEGFRMLSPBEHAPAEYDINKTCWDA 909  
QY 929 DSKRSPFNLTSLFGLCOLADAEAMYNQV 958  
DB 910 DPLKRPFTFKQIVQVLEKQISESTNHHYSNL 939



RESULT 8

IS1703  
C-kit-related kinase 1 (XKrk1) - African clawed frog  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 18-Jun-1999  
C:Accession: IS1703  
R:Baker, C.V.; Sharpe, C.R.; Torpey, N.P.; Heasman, J.; Wylie, C.C.  
Mech. Dev. 50, 217-228, 1995  
A:Title: A Xenopus c-kit-related receptor tyrosine kinase expressed in migrating stem cell  
A:Reference number: IS1703; MUID:95344996; PMID:7619732  
A:Accession: IS1703  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1-954 <BAK>  
A:Cross-references: EMBL:Z48770; NID:G763033; PIDN:CAA8688.1; PID:G763034  
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology  
C:Keywords: ATP  
F:575-915/Domain: protein kinase homology <KIN>  
F:583-591/Region: protein kinase ATP-binding motif

Query Match 23.2%; Score 1226; DB 2; Length 954;  
Best Local Similarity 33.1%; Pred. No. 1.5e-52;  
Matches 305; Conservative 162; Mismatches 322; Indels 132; Gaps 30;  
QY 107 FKHSILNCPHFRLQNRGVVS-----MVLKMTETQAGEYLLFIQSEATNYTILFTVSI 160  
DB 52 FQSGMLKKPR-DLKRPLNNSTDTQFVLIKADLRHIGYICTNTQENTSV--SLFV 108  
QY 161 RNTLLYLRPRFRKMNQDALVCSIESVPEVTEVWVLCDSQGESCKE-----ESPA 212  
DB 109 KDPARPFLLDPIFDVTEGADTVGCMCFPTDMDIAIEKCD--GSPLENFTFTDIEAGI 166  
QY 213 VVKKEKVLHLEFGTDIRCARNELGR-ECTRLFTIDLNQTPQTLLPOLFL-----KV 264  
DB 167 TIKTVQLAFDSCV-----VCSGNKSTGVKSTSTFHVKVPVK-KVPTVFLSKSRQLVKT 220  
QY 265 GEPMLTRCKAVNVHGFGLTW-ELNKALEEGNYFEMSTYSTNRTMIRLILFAPVSSVARN 323  
DB 221 GEPFVTCVAVLDVFTSVKQMLDVGVKTKQANFRSNVFSNLTLS-----DGVVYS 274  
QY 324 DTGYTCSSSKHPQS---SALVTIVGKFINAT--NSSEYEDIDYEEFCFSVRKAYPO- 378  
DB 275 ESRFTTCQENAIQGVNATFTLDVIDGVYNTLVLENTTISVNGDNLVVKVYIDAYPH 334  
QY 379 IRCTVTSKSPFCEOKGLDNGYSISK-----FCNHKH-----QPGYIFHAENDDA 425  
DB 335 DDGVWTFYNETL---LNTSDHYVATKDEGNRYVSELHLRLKGTGKGYTFYTTNSDD 390  
QY 426 QTKMFTLIRKPKQVLAASASQAS--CFSDGYPLPSWTWKCKSDKSPNCTEETEGVW 483  
DB 391 DASVSFNIQVTRPELIIAERTSEGLQCVATGFPVPAIQWFCPSGEORCTDYPLSPV 450  
QY 484 NRK--ANRNVFGQWSSSTLNMSSEATKGLFVKKCAVNSLGTSCETILLNSPGPFPE-IO 540  
DB 451 NEKFTQENSLGRIVVESTIDVNDLKNKGTVCVASNEVESAYSV-----FSPAIKE 502  
QY 541 N-----ISFYATIGVCLLFTVLVTLFLCHKYKQFRYESQLQW-QVTGSSDNE 588  
DB 503 KLRTHTLFTPLLIIGFTAAAGLMCIAVAVL---MYKYLQPKYEIQWVVEEING---NN 555  
QY 589 YFVDFPREYDYLKWFPPRENIIEFGKVLGSAFGKMNATAYGISITGVSIQVAVQMLKE 648  
DB 556 YVYIDPTQPLYNKWFPPRDLFCFGLGAGAGKVEATAYGLLKEDSRLLTAVRMLKP 615  
QY 649 KADSSERREALMSLKMOTGLSHENIVNLIGACTLSGPIYLLIFCYCCYGLLNLRSKEE 708  
DB 616 SAHSTEREALMSLKVSLYGLHHKNIVNLIGACTVGGPILVITEYCCYGLLNLRRKED 675  
QY 709 KPHRTWTEIFKEHNFSTFYTFOSHPS-----SMPGSRVQIHPDSQI----- 752  
DB 676 SF-----ICPKFEDNSAALYKKNLNTDMGCEGSEYIDMKPAVSVVVPT 721  
QY 753 ---SGLHGNSFHSEDEIEYENQKRLBEBDLNLTLPEDLLCFAYQVAKGMEFLFKSCVHR 810

DB 722 KTDKRSGSGFGQDQV-----SVSIPEDDL-ALDTEDLNFYSQVAGQNMFLASKNCIHR 775  
QY 811 DLAAARNVLVTHGKVKVCDPGLARDIMSDSNVYVGNARLPVKWMAPESLFEGIYTIKSD 870  
DB 776 DLAAARNILLTHGRITKICDFGLARDIRDSNVYVGNARLPVKWMAPESTFHCVTTFESD 835  
QY 871 VNSYGTLLWEIESLGNVPYGPIDANFYKLQNGPKMDOPFVATEEYIIMQSCWAFDS 930  
DB 836 VNSYGTLLWEIESLGSPPRIPVDSKFKYMKIKDGYRMSPECAPLEMVEIMRSCWNSDP 895  
QY 931 RXRPSPNLTSLPLGCOLADAE 951  
DB 896 LKRPFKQIVQVMEQQLSDSK 916

RESULT 9

TVCTMD  
macrophage colony-stimulating factor 1 receptor precursor - cat  
N:Contains: protein-tyrosine kinase (EC 2.7.1.112) csflr/fms  
C:Species: Felis silvestris catus (domestic cat)  
C:Date: 31-Dec-1989 #sequence\_revision 31-Dec-1989 #text\_change 13-Jun-1997  
C:Accession: A31636  
R:Woolford, J.; McAuliffe, A.; Rohrschneider, L.R.  
Cell 55, 965-977, 1988  
A:Title: Activation of the feline c-fms proto-oncogene: multiple alterations are required  
A:Reference number: A31636; MUID:89077553; PMID:2849512  
A:Accession: A31636  
A:Molecule type: mRNA  
A:Residues: 1-980 <WOO>  
A:Cross-references: EMBL:X03663  
C:Genetics:  
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;  
C:Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming protein;  
F:1-23/Domain: signal sequence #status predicted <SIG>  
F:24-980/Product: macrophage colony-stimulating factor 1 receptor #status predicted <EXT>  
F:24-509/Domain: extracellular #status predicted <EXT>  
F:35-86/Domain: immunoglobulin homology <IMM1>  
F:120-179/Domain: immunoglobulin homology <IMM2>  
F:217-280/Domain: immunoglobulin homology <IMM3>  
F:316-381/Domain: immunoglobulin homology <IMM4>  
F:410-484/Domain: immunoglobulin homology <IMM5>  
F:510-534/Domain: transmembrane #status predicted <TMM>  
F:535-980/Domain: intracellular #status predicted <INT>  
F:577-915/Domain: protein kinase homology <KIN>  
F:585-593/Region: protein kinase ATP-binding motif  
F:42-84,127-177,224-278,417-482/Disulfide bonds: #status predicted  
F:45,73,94,153,275,302,335,410,477,490/Binding site: carbohydrate (Asn) (covalent) #stat  
F:613,630,776/Active site: Lys, Glu, Asp #status predicted  
F:781,794/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 23.2%; Score 1225.5; DB 1; Length 980;  
Best Local Similarity 33.5%; Pred. No. 1.6e-52;  
Matches 321; Conservative 168; Mismatches 321; Indels 149; Gaps 34;  
QY 83 VDVASITLQVLVDAPGNISCLW---VFKHSILNCP-----HFDLQNRGVSMVIL 131  
DB 32 VEPGTTVLRV-----GNGSVWDGPISEHWNLDLDPSSILTNNTATQNTGYHCT-- 85  
QY 132 KMTETQAGEYL--LFTQSEATNYTIL---FTVSIRNTLLYLRPRFRKMNQDALV-CI 185  
DB 86 EPGNPGGNATILYVKDPAKPKVLAQEVTV-----LEGDALLPCL 128  
QY 186 SESVEPIVE--WVLCDSQGESCKE-----SP---AVVKEEKLHLEFGTDIRCARN 235  
DB 129 ---LTDPALEAGVSLVRVRGRPVLRQTNYSFSPHGFTIHKAKFIENHVY----QCSARV 181  
QY 236 ELGRECTIL-----FTIDLNQTPQTLLPOLFLKV-GEPLWIRCKAVNVHGFGLTWELE 288  
DB 182 D-GRVTSGIWLKVKQKDIISGPATITLLEPAELVRIQGEAAQIVCSASNDIVNPDV----- 235





Qy 836 IMSDSYVVRNARLPVKWAPESIFEGITYIKSDVWSYGILLWEIFSLGVNYPGPVD 895  
 Db 801 IMDSNYIVKGNARLPVKWAPESIFDCVYVQSDVWSYGILLWEIFSLGVNYPGILVN 860  
 Qy 896 ANFYKLIQNGFMQDPFYATBEIYIIMQSCWAFSRKPSFNLTSFYGCCLADAEAMY 955  
 Db 861 SKFYKLVKDYQMAQAPAPAKNIYSIMQCAWALEPRTREPTFOQICSLLOKQ---AQE--- 914  
 Qy 956 QNVGGRV---SECHTYQNR--RPF 975  
 Db 915 ---DRVPENYTNLPSSSSSRLRPW 936

RESULT 11  
 TVHUMD  
 N:Macrophage colony-stimulating factor 1 receptor precursor - human  
 C:Species: Homo sapiens (man)  
 C>Date: 28-Dec-1987 #sequence revision 31-Dec-1993 #text change 16-Jul-1999  
 C:Accession: S08123; A24533; I56672; I57648; I59083; I52772  
 R:Hampe, A.; Shamoon, B.W.; Gobet, M.; Sherr, C.J.; Galibert, P.  
 Oncogene Res. 4, 9-17, 1989  
 A:Title: Nucleotide sequence and structural organization of the human FMS proto-oncogene  
 A:Reference number: S08123; MUID:89239490; PMID:2524025  
 A:Accession: S08123  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-972 <HAM>  
 A:Cross-references: GB:U63963; EMBL:X14720; NID:g1915975; PIDN:AA851696.1; PID:g1915976  
 A:Note: this sequence was submitted to the EMBL Data Library, March 1989  
 R:Cousseens, L.; Van Beveren, C.; Smith, D.; Chen, E.; Mitchell, R.L.; Isacke, C.M.; Verm  
 Nature 320, 277-280, 1986  
 A:Title: Structural alteration of viral homologue of receptor proto-oncogene fms at carb  
 A:Reference number: A24533; MUID:86175013; PMID:2421165  
 A:Accession: A24533  
 A:Molecule type: mRNA  
 A:Residues: 1-53, 'A' 55-972 <COU>  
 A:Cross-references: GB:J03149  
 A:Note: the authors translated the codon GCA for residue 54 as Pro  
 R:Wheeler, E.F.; Roussel, M.F.; Hampe, A.; Walker, M.H.; Fried, V.A.; Look, A.T.; Retter  
 J. Virol. 59, 224-233, 1986  
 A:Title: The amino-terminal domain of the v-fms oncogene product includes a functional e  
 sequences.  
 A:Reference number: I56672; MUID:86281820; PMID:3525854  
 A:Accession: I56672  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-16 <RES>  
 A:Cross-references: GB:M14002; NID:g182676; PIDN:AAA35849.1; PID:g553292  
 R:Visvader, J.; Verma, I.M.  
 Mol. Cell. Biol. 9, 1336-1341, 1989  
 A:Title: Differential transcription of exon 1 of the human c-fms gene in placental troph  
 A:Reference number: I57648; MUID:89261741; PMID:2524648  
 A:Accession: I57648  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-16 <RE2>  
 A:Cross-references: GB:M25786; NID:g349454; PIDN:AAA58421.1; PID:g553224  
 R:Browning, P.J.; Bunn, H.F.; Cline, A.; Shuman, M.; Nienhuis, A.W.  
 Proc. Natl. Acad. Sci. U.S.A. 83, 7800-7804, 1986  
 A:Title: Replacement of COOH-terminal truncation of v-fms with c-fms sequences markedly  
 A:Reference number: I59083; MUID:87017034; PMID:3532121  
 A:Accession: I59083  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 874-972 <RE3>  
 R:Nienhuis, A.W.; Bunn, H.F.; Turner, P.H.; Gopal, T.V.; Nash, W.G.; O'Brien, S.  
 Cell 42, 421-428, 1985  
 A:Title: Expression of the human c-fms proto-oncogene in hematopoietic cells and its del  
 A:Reference number: I52772; MUID:85282599; PMID:4028159  
 A:Accession: I52772  
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
 A:Residues: 244-295 <RE4>  
 A:Cross-references: GB:M11067; NID:g182674; PIDN:AAA35848.1; PID:g442423  
 C:Genetics:  
 A:Gene: GDB:CSPLR; FMS  
 A:Cross-references: GDB:120600; OMIM:164770  
 A:Map position: 5q33.2-5q33.3  
 A:Introns: 17/1; 103/1; 198/1; 243/3; 297/1; 361/2; 400/1; 440/2; 504/1; 542/3; 585/1;  
 C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;  
 C:Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming protei  
 fic protein kinase  
 F:1-23/Domain: signal sequence #status predicted <SIG>  
 F:24-972/Product: macrophage colony-stimulating factor 1 receptor #status predicted <M  
 F:24-512/Domain: extracellular #status predicted <EXT>  
 F:35-86/Domain: immunoglobulin homology <IMM1>  
 F:120-179/Domain: immunoglobulin homology <IMM2>  
 F:217-280/Domain: immunoglobulin homology <IMM3>  
 F:316-383/Domain: immunoglobulin homology <IMM4>  
 F:412-487/Domain: immunoglobulin homology <IMM5>  
 F:513-537/Domain: transmembrane #status predicted <TM>  
 F:538-972/Domain: intracellular #status predicted <INT>  
 F:580-917/Domain: protein kinase homology <KIN>  
 F:588-596/Region: protein kinase ATP-binding motif  
 F:42-84, 127-177, 224-278, 419-485/Diulfide bonds: #status predicted  
 F:45, 73, 153, 240, 275, 302, 335, 353, 412, 428, 480/Binding site: carboxydrate (Asn) (covalent);  
 F:616, 633, 778/Active site: Lys, Glu, Asp #status predicted  
 F:783, 796/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 23.1%; Score 1216.5; DB 1; Length 972;  
 Best Local Similarity 32.8%; Pred. No. 4.4e-52;  
 Matches 325; Conservative 156; Mismatches 316; Indels 195; Gaps 33;

Qy 75 YVEAAVEVDV--SASITLQVLVDAPGNISCLWFKHSLNCQPHFDLQNRGVSNVILK 132  
 Db 22 VIEPSVPELVVPGATVTLRCV---GNGSVEMDGPSS---PHWTLYSDGSSILSTN 72  
 Qy 133 MTE-----TQAGEYL-----LFIQSEATNTTLFTVSIINTLLYTLRRYFRKM 176  
 Db 73 NATQNTGTGYCTSPGDLGSAIHLVKDPAHPNNVL---AOEVVVF----- 118  
 Qy 177 ENQDALV-CISESPEPEVPEVWLCDSCQESCKEESPAAVKKEEK-----VLHE--- 223  
 Db 119 EDQALLPCL---LTDPLVE-----AGSLVRVGRPLMRHTNYSFSPWHGFTTHRAKF 169  
 Qy 224 LFGTDIRCCARNELGRECTRLTIDLNOTPTTLP-----QLFLKVGPELWIRCK 273  
 Db 170 IOSODYQCSALMG-GR---KVMISIRLKVQKVPGLPPALTLVPAELVIRGEAAQIVCS 225  
 Qy 274 AVHYNHGFLTWELNKALEBGNFEMSTYTNRTMIRILFAFVSSVARNDTGYTCSSS 333  
 Db 226 ASSVDVNFDFVLOHNNTKLA---IPQSDPHNNRYQ-KVLTLNLDQVDFQHAGNYSVAS 281  
 Qy 334 ---KHPQSALVTIVGKGFNATNSED---YEIDQVEECFVSVPKAYPOIR-CTWTF 385  
 Db 282 NVQGRH-STSMFRRVRESAYLNL--SSEONLQIEVTVEGLNLKVMVEAYPGLQGFNWTY 338  
 Qy 386 SRKSFPCQKGLDNGYSISKFCNKHQP-----GEYTF 418  
 Db 339 -----LGFSDHQPEPKLANATKDYRHTFTLSLPLRKPSEAGRYSF 381  
 Qy 419 HAENDDAQTKMTLNRKPKQVLAESASQAS-----CFSDGYPLPSTWTKCKSDKSPNC 474  
 Db 382 LARNPGGWRALTFTLTLYRPPESVVIWTFINGSGTLLCAASGYPPQNPVTLQCSGHTDR 441  
 Qy 475 TEEITEGVN---RKANRKYVGQWSSSTLNMSEAIGFLVKCCAYNSLGTSCETILLN 530  
 Db 442 DEAOVLQWDDPYDEVLSQEPFHKVTYVOSLTLTETLEHNTQYECRAHNSVSGGSNAFPI 501  
 Qy 531 SPGPFPTQDNISFYATIGV---LLFTVVLTLTICHYKKQFRYESQLQVMQVVTGSSN 587  
 Db 502 SAGATHPPDEFLTTPVWVACMSIMALLLLLLLYKKQPKYQVRWKIIE--SYEGN 559  
 Qy 588 EYFYVDREYEDLKWEPRENLEBFGKVLGSGAFGKVMNATAYGISKTGVSITQVAKMLK 647

Db 560 SYTFIDPTQLPYNEKWEPPRNQLQKTLGAGAFKVVVEATAFGLGKEDAVLKVAVMMLK 619  
Qy 648 EKADSSEREALMSLKMWTOLGSHENIVNLGACTLSGPIYILIPFYCCYGLDNLNLSKR 707  
Db 620 STAHADKEALMSLKMWTOLGSHENIVNLGACTLSGPIYILIPFYCCYGLDNLNLSKR 679  
Qy 708 E-----KFHRTWTEIPKHNFPSPYPTFQSHNPSMPSGSRVQIHP 747  
Db 680 EAMLGPSLSPGQDPGGVDYKNIHLKRYVRDGSFS-----SQGVDTYV 724  
Qy 748 DSDQISGLHNSHSEDEIEYENQKRLKEEDLNVLTFEDLLCFAYQVAKGMEPLKPC 807  
Db 725 EMRPVSTSSNSF-SEQDLKEDGRLP-----LRDLHFSQVAGKAFKASNC 774  
Qy 808 VHRDLAARNVLVTHGKVKIKDFGLARDIMSDNSVVRGNARLPVKWMAPELSFEGIYTI 867  
Db 775 IHRDVAARNVLVTHGKVKIKDFGLARDIMSDNSVVRGNARLPVKWMAPELSFEGIYTI 834  
Qy 868 KSDVMSYGILLWEIFSLGVPNPGIPVDANFYKLQNGPKMDQPPYATEEIIIMOSQWA 927  
Db 835 QSDVMSYGILLWEIFSLGVPNPGIPVDANFYKLQNGPKMDQPPYATEEIIIMOSQWA 894  
Qy 928 FDSRKRPSPFNLTSLGCOL--ADAEEMQYV 958  
Db 895 LEPTHRPTFOQICSLQBOAQEDRRERYTNL 926

RESULT 12  
JN0677  
protein-tyrosine kinase (EC 2.7.1.112), receptor type kit precursor - chicken  
N;Alternate names: tyrosine kinase receptor kit  
A;Title: Cloning and expression of the chicken c-kit proto-oncogene.  
C;Species: Gallus gallus (chicken)  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Jun-2000  
C;Accession: JN0677  
R;Sasaki, E.; Okamura, H.; Chikamune, T.; Kanai, Y.; Watanabe, M.; Naito, M.; Sakurai, M.  
Gene 128, 257-261, 1993  
A;Title: Cloning and expression of the chicken c-kit proto-oncogene.  
A;Reference number: JN0677; MUID:93292995; PMID:7685729  
A;Accession: JN0677  
A;Molecule type: mRNA  
A;Residues: 1-960 <SAS>  
A;Cross-references: DBJ:D13225; NID:g303532; PID:BAA02506.1; PID:g303533  
A;Experimental source: brain  
C;Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;  
C;Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming protein;  
protein kinase  
F;1-24/Domain: signal sequence #status predicted <SIG>  
F;25-960/Product: tyrosine kinase receptor #status predicted <MAT>  
F;314-380/Domain: immunoglobulin homology <IMV>  
F;573-916/Domain: protein kinase homology <KIN>  
F;581-589/Region: protein kinase ATP-binding motif  
F;76,135,149,269,286,306,318,338,356,453,469/Binding site: carbohydrate (Asn) (covalent)

Query Match 23.0%; Score 1215; DB 1; Length 960;  
Best Local Similarity 32.3%; Pred. No. 5.1e-52;  
Matches 314; Conservative 165; Mismatches 370; Indels 124; Gaps 34;

Qy 66 ALRPOSSGTVYAAAEVDVSAITLOVLVDAPGNISCLVWFKHSS-----LNQPHFD 119  
Db 18 SLIPAGGVPHESSLVWNGKBEALKNEEGP-----VTWNFQNSDPSAKTRISNEKEWH 73  
Qy 120 LQNRGV--VSMVLKMTQOAGEYLLFIOSEATNYTLFTSVIRNLLYLRPFRKME 177  
Db 74 TKNATIRDIGRECKSGKSVNSFYVFKDP-----NVLFV---DSLIV-----GKED 119  
Qy 178 NODALVCSISEPPEP-IVEVNLCDSGESCKEESPAVVKKE-----KVLHFLFGTDIRCC 232  
Db 120 SDILLVC---PLTDPDLNFTLRKCDGKPLKXNMTPIPNQKGIILKVNQRFKGCYQCL 176  
Qy 233 AR-NELGRECTRLFTIDLNQTP-QTTLPOI-----FLKVGELPLWIRCKAVHNVHFG 283  
Db 177 AKHNGVEKISEHIF---LNVRPVHKALPVITLSKSYELLKEGEEFEVTCITIDVDSSVKA 233

Qy 284 TWELNKALEGNYPEMSTYSTNRTMIRILFAPYSSVARNDTGYVTCSSKHP--SQSAL 341  
Db 234 SWISYKSAIVTSKSRNIGDYGERK-----LTLNIRSVGNDSGEFTC-QAENPFQKTNAT 288  
Qy 342 VTI--VGKGFN--ATNSSDEYIDQVEEFCFSVRPKAYQIIR-CTWTFGRKSPCEQK 395  
Db 289 VTKALAGFVRLPATWNTTIDINAGQNG--LIVEYEAYPKPEEVWYNNETL--QN 343  
Qy 396 GLDNGYSIKFCNKKH-----OPEYIFHAENDDAQFTKMTFLNIRRPQVLA 443  
Db 344 SSDHVVPKTVGNNSYTSSELHLTKGTEGGIYTPFFVNSDASSSVTVNVVTKPEILT 403  
Qy 444 EASASQ--ASCFSDGYPPLPSWTWKCKSDKSPNCTEETEGVWNRKANRKFQWVSSSTL 501  
Db 404 LDMGLNDILOCVATGFPAPTYYVPCGTQRCLODSTISPMQVKS-----YTNSSVP 457  
Qy 502 NMSEAIKGLFKCCAYNSLGTSCETILLNSPGPPF-----IQDNISFYA-----TIGV 550  
Db 458 SPERILVESTVNASMFKSTGTCCEASNGDKSSVFFNFAIKEQIRHTLPTPLLIAGV 517  
Qy 551 CLLFIVVLTLICHKKYKQFYESOLOMV-QVTGSSDNEYFYVDPREYEDLKWFPFRN 609  
Db 518 AAGLWCIIIMILVIYIYLOPKYEQVQVVEING--NNYVYIDPTQLPYDHKWEFPRN 574  
Qy 610 LEFGKVLGSGAFGKVMNATAYIGISKTGVSIOAVVOMLKEKADSSREALMSLKMQLG 669  
Db 575 LSPGKTLGAGAFGKVVVEATAYGLFKSDAAMTAVAVKMLKPSAHLTREALMSLKVLSYL 634  
Qy 670 SHENIVNLGACTLSGPIYILIPFYCCYGLDNLNLSKRKEP---HRTWTEIFKEHNPSF 725  
Db 635 NHINIVNLGACTIGGPTLVITEYCCYGLDNLNLRKRDSPICPKHEEHAEEAVENL-- 692  
Qy 726 YPTQSHSPNS-----SMPGSRREYQIHPDSDQISGLHNSHSEDEIEYENQKLEEE 778  
Db 693 --LHQAEPTADAVNEYMDKPGVSYAVPPKADKKPKVSGSYTDQD---VTLSMLEDD 746  
Qy 779 DLNVLTFEDLLCFAYQVAKGMEPLKFCVHRDLAARNVLVTHGKVKIKDFGLARDIMS 838  
Db 747 --LALDVEDLLSFSYQVAKGMSFLASKNCIHRDLAARNILLTHGRITKICDFGLARDIN 804  
Qy 839 DSNVYVRGNARLPVKWMAPELSFEGIYTIKSDVMSYGILLWEIFSLGVPNPGIPVDANF 898  
Db 805 DSNVYVRGNARLPVKWMAPELSIFNCTYTFSDVMSYGILLWEIFSLGSGSPYPMGVDSKF 864  
Qy 899 YKLIQNGFKMDQPPYATEEIIIMOSWAFPSRKPSFPNLTSLFLGCOLADAEEMQYV 958  
Db 865 YKMIKEGYMFSPECSPEMYDIMSCKWDADPLQRTPKQIVQLIEQQLSDNAPRYAN- 923  
Qy 959 DGRVSECPHTYON 971  
Db 924 --FSTPPSTQGN 933

RESULT 13  
T30816  
macrophage colony-stimulating factor receptor - Japanese pufferfish  
C;Species: Fugu rubripes (Japanese pufferfish)  
C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 02-Jun-2000  
C;Accession: T30816  
R;How, G.F.; Venkatesh, B.; Brenner, S.  
Genome Res. 6, 1185-1191, 1996  
A;Title: Conserved linkage between the pufferfish (Fugu rubripes) and human genes for p.  
A;Residues: 1-975 <HOW>  
A;Accession: T30816  
A;Status: preliminary  
A;Molecule type: DNA  
A;Cross-references: EMBL:U63926; NID:g1752706; PID:g1752708; PID:AA60063.1  
C;Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;  
Query Match 23.0%; Score 1213; DB 2; Length 975;  
Best Local Similarity 32.4%; Pred. No. 6.5e-52;

Matches	320;	Conservative	180;	Mismatches	347;	Indels	140;	Gaps	31;
QY	75	VYEAAAEVDVSASITQLQVLDPAGNISCLW	--VFKESSLNCQHPDLQNRGVVSMVL	131					
Db	31	VVQSEVVVKPGTSLLELCKGGDPVN	---WQTRLPKHKRYMSRSPGNLR	---	TRVA	81			
QY	132	KMTETQAGEYLLFIQSEATNYILFTVSI	-----RNTLLYTLRRPYFRMENQD	180					
Db	82	RLTAEFTGYKCFYGAWAQHRHLTSSVHYVKDPNRVFWTSTSLRVVRK	-----RGED	135					
QY	181	ALVCISESPEPIVEWVLCDGCEKKEESPAVVKEBKV	---LHELFGTDIRCCARNE	236					
Db	136	YLLPCLLTDPETDLGLRNDN-GTIVPEPMNYTVYVRHGILIRSLQPSFNADYVCTAKVK	194						
QY	237	LGRECTRLFTIDLNQTPQTTLPQLFLK	-----VGEPLWIRCKAVHVNHGFGLTWELEN	289					
Db	195	GVKEKTSKTFINVIQKLRFPP-PYVPLEMDEVYRIVGEBELQIRCTHNPENFNVTWNTVT	253						
QY	290	KA--LEEGNYFEMTSTYNTNRIMILAFVSSVARNDTGYTTCSSSKHP	---SQSALVT	343					
Db	254	KSRVTIEE---RVKSSGENRLDQSIULT-ISAVDLADTGNISCICTNEAGVNSNTYLL	308						
QY	344	IVGKGF1-----NATNSSEDIYIDQYEEFCFSVRFKAYPOIR-CTWTFPSRKSPFCEQ	394						
Db	309	VVEKPYIRLWPOLIPKLASQGLSVEVNEGDELGLVMVEAYPQITDRHWHTPSPSTSMQ	368						
QY	395	KGLONGYSISFECNKHQPGEYIFHAENDDAQFTMTFLNIRRK-PQVLAEASASQASCF	453						
Db	369	EHIYHARLQLKRMNAQEQ-GOYTFYAKNSLANGSISFHVVMYQKPIAVRWENITTLTCT	427						
QY	454	SDGYLPSTWTKKCDKSPNCTEBITEGVNWR	-----KANRKVFGQWSSSTLNWSE	505					
Db	428	SFGYPAQIILWYQCSGIRPTCNGNT-GLPKQNIHQALTEVQREEGAVEVESVFTVGL	486						
QY	506	AIKGLVKCCAYNSLGTSCETILNSPGPPFIQDNISFYATI	----GVCLLFTIVLITLL	561					
Db	487	SNHRMTVECAFLNVGVSSDFTTVE	-----VSDKL-FTSTLIGAAGVLAI	538					
QY	562	ICHVKYKQPRVESQLQWQVVTGSSDNEYFYVDREYDYDLKWFPRENLEFQKVLGSGAP	621						
Db	539	L-YIKQKQPRFEIRKWKIE--AREGNNTYTFIDTQLPYNEKWEFFPRDKLGLKVLGAGAF	595						
QY	622	GKVNNTAYAGISK-TGVSIOVAVKMLKEKADSSEREALMSLKMOTOLGSHENIVNLGA	680						
Db	596	GKVVEATAFGLGEDKNTLRVAVKMLKANAHSDEREALMSLKLHSLGHQHQIVNLGA	655						
QY	681	CTLSGPIYLPYCCYGDLLNLYLSKREKFKHRTWTEIFKBNHNSFYPTFQSHPN	-----734						
Db	656	CTYGGPVLVITEYCSLGDLLNFLRQKAETFVNLVWNI	-----PEIMENSNDYKNIC	706					
QY	735	-----SSMPGSRREVOIHPDSDQISGLHGNSFHSEDEIYENQKRLLEEEDLVN	782						
Db	707	NQKWIIRSDGISISTSSSTYLEMRPS	-----QOQSHIEASGRKSLCEDNGDW	753					
QY	783	LTPEDLLCFAYQVAKGMFELEFKS CVHRDLAARNVLVTHGKVVKI	CDFGIARDIMSDSNY	842					
Db	754	LDIDDLRLFSLQAQGLDFLASRNCIRHDAARNVLLTDKRVAKI	CDFGIARDIMSDSNY	813					
QY	843	VVRGNARLPVKWMAPESLPGIITIKSDVWSYIGILLWEI	FPISGLGNVPYPGIPVDANPYKLI	902					
Db	814	VVRGNARLPVKWMAPESI	IFPCVTVQSDVWSYIGILLWEI	PSLGKSPSPMAVDSRYKMW	873				
QY	903	QNGFKMDQPPYATEEYIIMQSCWAPDSRKRPFPNLT	TSFLGGQLA--DAEAMVQNV-	958					
Db	874	KRGVQMSQDPFALPEIYIMIMKMCWNLEPTERTPTFSMI	SQMINELLCGQDEQEKLIYRNQ	933					
QY	959	-----DGRVSECHTYQNRPPPREMD	980						
Db	934	PEQVAEGEACDEPKRYD--PPCERSCD	958						

RESULT 14  
S16385

[illegible]

Query Match 22.6%; Score 1194; DB 2; Length 978;  
Best Local Similarity 33.1%; Pred. No. 5.5e-51;

Qy	83	VDVSASITLQVL-----YDAPCNISCLMVFVKHSSLNCPHFQDQNRGVV-----SMVLTKM	133
Db	32	VEPGETITLRCVNGSVEMDGPISPYWTLDPES-----PGSTLTTRNATFKNTGYRCTEL	87
Qy	134	TETQAGYL--LFIQSEATNYTL--FTVISRNTLLYLRPFYRKEMODALVCISES	188
Db	88	EDPMAGSTTHLYVKDPAHWSNLLAQEVTV-----VEGEAV-----	124
Qy	189	VPEPIVEMWLCDSGESKEESPAVVK-----BKVLHELPGTDIRCCAR	234
Db	125	LPCLITDPAUKDSV-SLMREGGRQLRKTYFFSAMRGFIIRKAKVL-----DSNTYVCKT	179
Qy	235	NELGRECTRL-FTIDLN---QTPQTTL-PQLFLKV-GEPLMTRCKAVHVNHGSLTWEL	287
Db	180	MVNGRESTGIVLKNRVHPEPQIKLEPSKULVRIRGEAAQVCSATNAEVGNFVILKR	239
Qy	288	ENKALE-----EGNYFEMSTYSTNRMTIRILFAFVSVARNDGYTTC-SSSKHPQS	339
Db	240	GDTKLEIPLNSDFQDNYYK-----KVRLSINAVDQAGIYCSVANDEVGTRT	288
Qy	340	ALVT--IVKGFINATN--SEDEYIDQYBEFCFSVRFKAYPQIR-CTWTFSRKSFPCEQK	395
Db	289	ATMNFQWESAVILNLTSEQSLQLEVSQVGSLSILTVHADAYPSIOHYNWTYLGPPFE--DQR	347
Qy	396	GLD-----NGYSISKFCNH--KHQPGEYI PHAENDDAQFTKMPTILNTRRKQVLAEAS	446
Db	348	KLEFITQIRAIRYTFKFLFLNRVKASAGGYFLMAQNKAGMNNLTFLTLRYPPEVSVTWM	407
Qy	447	ASOAS-----CFSDGYPLPSWTWKKCSKSPNCTEEITEGVVN-----RKANRKVFGQWVSS	498
Db	408	PVNGSDVLCFDYSGVPQPSVTWMECHGHTDRDCEAQLQVMDNTHPEVUSQKPEFDKVIQ	467
Qy	499	STLNMSBAIKGLVKKCAVNSLGTSCETILLNSPGFPFIQNISFYATIGVC---LLFI	555

```

Db      468  SOLPTGTLKHNMTYFKTHNSVNSQYFRVSLGQSKQIPDESLETPPVVACMSVMSLL 527
QY      556  VVLTLLCHKRYKQRYESQLOQVQVTSNDNEYFYVDYFREYEDLKWEPFRENLEFGKV 615
Db      528  VLLLLLLLYKQKPKYQVRWKIIE--RYEGNSYTFIDPTQLPYNEKWEFPNNLQFGKT 585
QY      616  LGSAGFGKVMNATAGISKTGVSIVQAVOMLKEKADSSEREALESKMTQOLGSHENIV 675
Db      586  LGAGAFGKVEATAFGLGKEDAVLKVAVOMLKEKADSSEREALESKMTQOLGSHENIV 645
QY      676  NLLGACTLGGPIYLIFCYCGDGLNLSRKE-----KEHRTWT 715
Db      646  NLLGACTHGGPVLVITEYCYCGDGLNLSRKEAMLGPSLPGQDSEGSSYNHLEKK 705
QY      716  EIFKEHNFSEYPTFOSHPSNPMPSREVQIHPDSQISGLHGNFSHSEDEIEYENOKRLE 775
Db      706  VYRRDSGFS-----SQVDYVEMRPVSTSSDSFPKQD-LDKEPSRPLE 749
QY      776  EEDLNVLTFEDLLCFAYOVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARD 835
Db      750  -----LWDLHFSQVAGMAFLASKNCITHRDVAARNVLLTSGHVAKIGDFGLARD 800
QY      836  INSDSNVYVRGNARLPVKWMAPESEFEGIYTKSDVMSYGILLWEIFSLGVNPPYGPVD 895
Db      801  INSDSNVYVRGNARLPVKWMAPESEFEGIYTKSDVMSYGILLWEIFSLGVNPPYGPVD 860
QY      896  ANFYKLIQNGFKMDQPFYATEETIYIMQSCWAPDSRKRPSFNLTSFLGCOLADAEAMY 955
Db      861  NKFYKLKVDGYOMAQPVPFAPKNYISIMQSCWDLPTFRPTFQOICFL-----QOARL 914
QY      956  QNVDRGVSECP 966
Db      915  ERRDQYANLP 925

RESULT 15
TWNSMD
macrophage colony-stimulating factor 1 receptor precursor - mouse
N:Contains: protein-tyrosine kinase (EC 2.7.1.112) csf1r/fms
C:Species: Mus musculus (house mouse)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 02-Jun-2000
C:Accession: S01880
R:Rothenwell, V.M.; Rohrschneider, L.R.
Oncogene Res. 1, 311-324, 1987
A:Title: Murine c-fms cDNA: cloning, sequence analysis and retroviral expression.
A:Reference number: S01880; MUID:88217329; PMID:2966922
A:Accession: S01880
A:Molecule type: mRNA
A:Residues: 1-976 <RQT>
A:Cross-references: EMBL:X06368
C:Genetics:
C:Gene: fms
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;
C:Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming protein;
fic protein kinase
F:1-19/Domain: signal sequence #status:predicted <SIG>
F:20-976/Product: macrophage colony-stimulating factor
F:20-515/Domain: extracellular #status:predicted <EXT>
F:35-86/Domain: immunoglobulin homology <IMM1>
F:120-179/Domain: immunoglobulin homology <IMM2>
F:217-280/Domain: immunoglobulin homology <IMM3>
F:316-381/Domain: immunoglobulin homology <IMM4>
F:410-485/Domain: immunoglobulin homology <IMM5>
F:516-535/Domain: transmembrane #status:predicted <TM>
F:536-976/Domain: intracellular #status:predicted <INT>
F:578-914/Domain: protein kinase ATP-binding motif
F:42-84,127-177,224-278,417-483/disulfide bonds: #status:predicted
F:45,73,302,335,389,410,449,478,491/Binding site: carbohydrate (Asn) (covalent) #status
F:514,631,776/Active site: Lys, Glu, Asp #status:predicted
F:781,794/Binding site: magnesium (Asn, Asp) #status:predicted

```

```

Query Match      22.4%; Score 1183.5; DB 1; Length 976;
Best Local Similarity 33.2%; Pred No.1.8e-50;
Matches 318; Conservative 154; Mismatches 347; Indels 139; Gaps 30;

QY      83  VDVSASITLOVL-----VDAPGNISCLWVF---KHSLNCQPHFDLQNRGVMSVILKMT 135
Db      32  VEPGETVTLRLCVNSGVSEWDGPIPIWTLDPSPGSTLTTSNATPKNTGYRCTELEDP 91
QY      136  TOAGBYLLFIQSBATNYTIL-----FTVSTRNTLYTLRRPYFRKMNQDALVCISEVPEP 192
Db      92  AGSTTHLYVKDPAHNSWLLAQEVTV-----VEQEAV-----LPCL 128
QY      193  IVEWVLCDSQSGSCKEESPAVVVKEEKVLHELFGTDIR-----CCARNELRECT 242
Db      129  ITDPALKDSV-SLMREGGRQVLRKTVYFSPWRGSIIRKAKVLDSTYVCKTMVNGREST 187
QY      243  RL-FTIDLN-----QTPQTL-LPOLFKV-GEPLWIRCKAVHVNHGFLGTLWELENKALE-- 293
Db      188  STGIWLKVNVRHPEPQIKLEPSKLVRIRGEAAQIVCSATNAEVRGNVILKRGDTKLEIP 247
QY      294  -----EGNYFEMSTYSTNRTMIRILFAFVSSVARNDTGYYTC-SSSKHPSQSALVT--IV 345
Db      248  LNSDPQDNYYK-----KVRALSLNAVDFQDAGIYSCVASNDVGTTRTATMNFQVV 296
QY      346  GKGFINATN-SSDEYIEDOYEFECFVRPKAYQPIR-CTMTFSRKSPFCEQKCLD----- 398
Db      297  ESAYNLNLTSEQLLOEVSVDGLIITVHADAPSIQHNWNTYLGPPFE-DQRKLEFITOR 355
QY      399  --NGVYSISKFNH--KHOPGEYIFHAENDDAQFTKMTLIRKPKOVLAESAASQAS--- 451
Db      356  AIYRTFKLFLNRVKASEAGQYFLAQNKAQNNLTFLTLYPPEVSVTWMPVNGSDVL 415
QY      452  -CFSDGYPLPSWTWKCKSDKSPNCTEETITEGVVN-----RKANRKFVGQWSSSTLAMEA 506
Db      416  FCDVSGYQPSVYTWMECRGHTDRCDQAALHLWNDTHPEVLSQKPPDKVIOISQLPIGL 475
QY      507  IKGFLVKCAVNSLGTSCETILLNSGPPFPFTQDNISFYATIGVC---LLFVIVLTLLIC 563
Db      476  KRMVYFCKTHNSVNSQYFRVSLGQSKQIPDESLETPPVVACMSVMSLLVLLLLLL 535
QY      564  HKYKQKRYESQLOQVQVTSNDNEYFYVDYFREYEDLKWEPFRENLEFGKVLSGAFGK 623
Db      536  YKYKQKPKYQVRWKIIE--RYEGNSYTFIDPTQLPYNEKWEFPNNLQFGKTLAGAFGK 593
QY      624  VNNATAYGISKTGVSIVQAVKMLKEKADSSEREALESKMTQOLGSHENIVNLLGACTL 683
Db      594  VVEATAFGLGKEDAVLKVAVKMLKSTAHADKEALMSKIMSHLGCHENIVNLLGACTH 653
QY      684  SGPIYLIFCYCGDGLNLSRKEKFRHTWTETIPEKHNFSPYPTFQSHPSNSMP----- 738
Db      654  GGFVLVYTYCYCGDHLNLRKAEAMHGP-----SLSPGQDSGDSYKNHLE 703
QY      739  -----GSGREVIQHPDSQISGLHGNFSHSEDEIEYENOKRLEEBEDNLVTFEDL 788
Db      704  KKYVRDSGFSQGVDTYVEMRPVSTSSDSFPKQD-LOKEHSRPLE-----LWDL 753
QY      789  LCFAYOVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDSNVYVRGNA 848
Db      754  LHFSSQVAGMAFLASKNCIHRDVAARNVLLTSGHVAKIGDFGLARDIMSDSNVYVRGNA 813
QY      849  RLPVKWMAPESEFEGIYTKSDVMSYGILLWEIFSLGVNPPYGPIDVANDFYKLIQNGFKM 908
Db      814  -LPVKWMAPESEFDCVITVQSDVMSYGILLWEIFSLGLNPNYPCGIHVNKNFYKLVKDYQ 872
QY      909  DDPFYATEETIYIMQSCWAPDSRKRPSFNLTSFLGCOLADAEAMYQNVDRGVSECP 966
Db      873  AQPVFAPKNYISIMQSCWDLPTFRPTFQOICFL-----QOARLERRDQYANLP 924

Search completed: August 26, 2003, 07:29:04
Job time : 37.0166 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2003, 07:14:08 ; Search time 19.0096 Seconds  
(without alignments)  
2456.527 Million cell updates/sec

Title: US-09-919-408A-4

Perfect score: 5274

Sequence: 1 MPALARDAGTVPLLVFSAM.....PFSRENDLGLSPQAQVEDS 993

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Swissprot\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5274	100.0	993	1	FLT3 HUMAN
2	4429.5	84.0	992	1	FLT3_MOUSE
3	1295.5	24.6	978	1	KIT_CAPHI
4	1286	24.4	977	1	KIT_BOVIN
5	1266	24.0	975	1	KIT_MOUSE
6	1251	23.7	976	1	KIT_HUMAN
7	1247.5	23.7	975	1	KIT_CANFA
8	1225.5	23.2	980	1	KFMS_FELCA
9	1224.5	23.2	978	1	KIT_FELCA
10	1224	23.2	978	1	KFMS_FSVMD
11	1216.5	23.1	972	1	KFMS_HUMAN
12	1215	23.0	960	1	KIT_CHICK
13	1204	22.8	977	1	KFMS_MOUSE
14	1194	22.6	978	1	KFMS_RAT
15	1176	22.3	1088	1	PGDS_RAT
16	1169	22.2	1089	1	PGDS_MOUSE
17	1166	22.1	1089	1	PGDS_HUMAN
18	1128	21.4	1087	1	PGDS_XENLA
19	1105.5	21.0	1098	1	PGDR_MOUSE
20	1086.5	20.6	1106	1	PGDR_HUMAN
21	990	18.8	1338	1	VGR1_HUMAN
22	989.5	18.8	370	1	KIT_FSVHZ
23	959.5	18.2	1333	1	VGR1_MOUSE
24	956	18.1	1336	1	VGR1_RAT
25	951.5	18.0	1356	1	VGR2_HUMAN
26	935	17.7	1348	1	VGR2_COTIJA
27	923	17.5	1367	1	VGR2_MOUSE
28	917.5	17.4	1343	1	VGR2_RAT
29	907.5	17.2	1363	1	VGR3_MOUSE
30	904.5	17.2	1298	1	VGR3_HUMAN
31	874.5	16.6	823	1	CEK3_CHICK
32	849	16.1	821	1	FGR2_HUMAN
33	845	16.0	813	1	FGR2_XENLA

34	827	15.7	821	1	FGR2_MOUSE
35	807	15.3	806	1	CEX2_CHICK
36	803.5	15.2	819	1	FGR1_CHICK
37	790.5	15.0	654	1	BFR2_HUMAN
38	784	14.9	812	1	FGR1_XENLA
39	779.5	14.8	822	1	FGR1_HUMAN
40	779	14.8	822	1	FGR1_MOUSE
41	766.5	14.5	806	1	FGR3_HUMAN
42	765	14.5	822	1	FGR1_RAT
43	754	14.3	801	1	FGR3_MOUSE
44	728.5	13.8	1052	1	FGR2_DROME
45	716	13.6	1115	1	RET_MOUSE

## ALIGNMENTS

### RESULT 1

ID	FLT3 HUMAN	STANDARD;	PRT;	993 AA.
AC	P36888; Q13414;			
DT	01-JUN-1994 (Rel. 29, Created)			
DT	01-JUN-1994 (Rel. 29, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	FL cytokine receptor precursor (EC 2.7.1.112) (Tyrosine-protein kinase receptor FLT3) (Stem cell tyrosine kinase 1) (STK-1) (CD135 antigen).			
GN	FLT3 OR STK1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=94119906; PubMed=7507245;			
RA	Small D., Levenstein M., Kim E., Carow C., Amin S., Rockwell P.,			
RA	Witte L., Burrow C., Ratajczak M. Z., Gewirtz A.M., Civin C.I.,			
RT	"STK-1, the human homolog of Flk-2/Flt-3, is selectively expressed in			
RT	CD34+ human bone marrow cells and is involved in the proliferation of			
RT	early progenitor/stem cells."			
RL	Proc. Natl. Acad. Sci. U.S.A. 91:459-463(1994).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93357464; PubMed=8394751;			
RA	Rosnet O., Schiff C., Pebusque M.J., Marchetto S., Tonnelle C.,			
RA	Toiron Y., Birg F., Birnbaum D.;			
RT	"Human FLT3/FLK2 gene: cDNA cloning and expression in hematopoietic			
RT	cells."			
RL	Blood 82:1110-1119(1993).			
RN	[3]			
RP	SEQUENCE OF 783-942 FROM N.A.			
RC	TISSUE=Testis;			
RX	MEDLINE=91169547; PubMed=2004790;			
RA	Rosnet O., Mattei M.-G., Marchetto S., Birnbaum D.;			
RT	"Isolation and chromosomal localization of a novel FMS-like tyrosine			
RT	kinase gene."			
RL	Genomics 9:380-385(1991).			
CC	- FUNCTION: RECEPTOR FOR THE FL CYTOKINE. HAS A TYROSINE-PROTEIN			
CC	KINASE ACTIVITY.			
CC	- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein			
CC	tyrosine phosphate.			
CC	- SUBCELLULAR LOCATION: Type I membrane protein.			
CC	- TISSUE SPECIFICITY: BONE MARROW CELLS.			
CC	- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-			
CC	PROTEIN KINASES			
CC	- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.			
CC	- DATABASE: NAME=PROW; NOTE=CD guide CD135 entry;			
CC	WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd135.htm".			
CC	- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;			
CC	WWW="http://www.infobiogen.fr/services/chronocancer/Genes/FLT3ID144.html".			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its



use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC	EMBL	U02687	AAA18947.1	--	
DR	EMBL	Z2652	CAA81393.1	--	
DR	EMBL	L36162	AAA35487.1	--	
DR	PIR	A39061	A39061	--	
DR	HSSP	P11362	1FGK	--	
DR	Genew	HGNC:3765	FLT3	--	
DR	MIM	136351	--	--	
DR	GO	GO:0005887	C:integral to plasma membrane	TAS	
DR	GO	GO:0005021	F:vascular endothelial growth factor receptor	--	TAS.
DR	GO	GO:0008284	P:positive regulation of cell proliferation	TAS	
DR	GO	GO:0007169	P:transmembrane receptor protein tyrosine kin.	--	TAS.
DR	InterPro	IPR007110	Ig-like	--	
DR	InterPro	IPR003599	Ig	--	
DR	InterPro	IPR003006	Ig_MHC	--	
DR	InterPro	IPR000719	Prot_kinase	--	
DR	InterPro	IPR001824	RTKinaseIII	--	
DR	InterPro	IPR001245	Tyr_pkinase	--	
DR	Pfam	PF00047	Ig; 1	--	
DR	Pfam	PF00069	pkinase; 1	--	
DR	ProDom	PD000001	Prot_kinase; 2	--	
DR	SMART	SM00409	IG; 1	--	
DR	SMART	SM00219	TyrcK; 1	--	
DR	PROSITE	PS50835	IG_LIKE; 1	--	
DR	PROSITE	PS00107	PROTEIN KINASE ATP; 1	--	
DR	PROSITE	PS50011	PROTEIN KINASE DOM; 1	--	
DR	PROSITE	PS00109	PROTEIN KINASE TYR; 1	--	
DR	PROSITE	PS00240	RECEPTOR TYR_KIN_III; 1	--	
DR	Signal	Transferrase	Tyrosine-protein kinase; Receptor; Transmembrane;	--	
KW	Glycoprotein	Phosphorylation	ATP-binding; Immunoglobulin domain.	--	
KW	Signal	1	POTENTIAL.	--	
FT	CHAIN	27	993	--	
FT	DOMAIN	27	543	--	FL CYTOKINE RECEPTOR.
FT	TRANSMEM	544	563	--	EXTRACELLULAR (POTENTIAL).
FT	DOMAIN	564	993	--	POTENTIAL.
FT	DOMAIN	253	343	--	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	610	943	--	IG-LIKE C2-TYPE.
FT	NP_BIND	616	624	--	PROTEIN KINASE.
FT	BINDING	644	644	--	ATP (BY SIMILARITY).
FT	ACT_SITE	811	811	--	BY SIMILARITY.
FT	CARBOHYD	43	43	--	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	100	100	--	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	151	151	--	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	306	306	--	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	323	323	--	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	351	351	--	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	354	354	--	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	473	473	--	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	502	502	--	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	541	541	--	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	8	8	--	A -> G (IN REF. 2).
FT	CONFLICT	10	11	--	TV -> QL (IN REF. 2).
FT	CONFLICT	78	78	--	A -> R (IN REF. 2).
FT	CONFLICT	227	227	--	T -> M (IN REF. 2).
FT	CONFLICT	346	346	--	G -> E (IN REF. 2).
FT	CONFLICT	940	940	--	T -> H (IN REF. 3).
SQ	SEQUENCE	993	AA; 112804	--	MM; 16790124B02F6BFB CRC64;

	Query Match	100.0%;	Score 5274;	DB 1;	Length 993;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 993;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MPALARDAGTVP	LLVFSAMIFGTTNQDLPV	IKCVLINHKNDSSVGKSSSPMV	SESP 60
Db	1	MPALARDAGTVP	LLVFSAMIFGTTNQDLPV	IKCVLINHKNDSSVGKSSSPMV	SESP 60
QY	61	EDLGCALRPOSSGTV	EEAAAVEVDVASITLOVLVDAP	GNISCLWVPKHSLLNCOPHFDL	120
Db	61	EDLGCALRPOSSGTV	EEAAAVEVDVASITLOVLVDAP	GNISCLWVPKHSLLNCOPHFDL	120

RESULT 2	FLT3 MOUSE	STANDARD;	PRT;	992 AA.
ID	FLT3 MOUSE			
AC	Q00342;			
DT	01-APR-1993	(Rel. 25, Created)		
DT	01-APR-1993	(Rel. 25, Last sequence update)		
DT	15-SEP-2003	(Rel. 42, Last annotation update)		
DE	FL cytokine receptor precursor (EC 2.7.1.112)	(Tyrosine-protein kinase		
DE	receptor flk-2) (Fetal liver kinase 2)	(Tyrosine-protein kinase FLT3).		
GN	FLT3 OR FLT-3 OR FLK-2.			
OS	Mus musculus (Mouse).			





```

Db 661 SELKMTLGHNDVNLGACTUSGVPYLIFCYCCGDLNLYRSKREPHRTWTIEFK 720
Qy 720 EHNFSYPTTQSHPSNMPGSRREVQIHPSDSQISGLHNSPHSDEIEYENQKEL--EEE 777
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 721 EHNFSYPTTQAHNSNMPGSRREVQVLPDQSGFNGSHSDEIEYENQKELAESEE 780
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 778 EDNLVLTFFELLCFAYQVAKGMEFLPKSCVHRDLAARNVLVTHGKVVVKICDFGLARDIM 837
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 781 EDNLVLTFFELLCFAYQVAKGMEFLPKSCVHRDLAARNVLVTHGKVVVKICDFGLARDIL 840
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 838 SDSNYVVRGNARLPVKWAPESLPEGYITKSDVMSYGILLWEIFSLGVNYPGIPVDAN 897
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 841 SDSYVVRGNARLPVKWAPESLPEGYITKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 898 FYKLIQNGFMDOPFYATEIYIIMQSCWAFDSKRPNSLTSFLGCLQCLADAEAMYQN 957
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 901 FYKLIQSGFMEQFPYATEGIYFWQSCWAFDSKRPNSLTSFLGCLQCLAEAEAC--- 957
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 958 VDRGVSECPHTYQNRPFPSREMDLGLLSPQAQVE 991
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 958 ----IRTSIHLPKQAAPQORG-GLRAQSPQORVK 986
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 3
KIT CAPHI
ID KIT CAPHI STANDARD; PRT; 978 AA.
AC Q28317;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)
DE (Proto-oncogene tyrosine-protein kinase Kit) (c-kit).
GN KIT.
OS Capra hircus (Goat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxID=9925;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Shiba; TISSUE=Cerebellum;
RX MEDLINE=97342548; PubMed=9199245;
RA Tanaka S., Yanagisawa N., Tojo H., Kim Y.-J., Tsujimura T.,
RA Kitamura Y., Sawasaki T., Tachi C.;
RT "Molecular cloning of cDNA encoding the c-kit receptor of Shiba goats
RT and a novel alanine insertion specific to goats and sheep in the
RT kinase insert region.";
RL Biochim. Biophys. Acta 1352:151-155(1997).
CC -!- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
CC 3-KINASE (PI3K).
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; D45168; BAA08116.1; --
CC HSSP; P11362; IFGK.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003599; Ig.

```

```

DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR001824; RTKinaseIII.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 2.
DR SMART; SM00409; IG; 3.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00835; IG_LIKE; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
KW Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
KW Immunoglobulin domain.
FT SIGNAL 1 22
FT CHAIN 23 978
FT DOMAIN 23 521
FT TRANSMEM 522 544
FT DOMAIN 545 978
FT DOMAIN 590 939
FT NE_BIND 596 604
FT BINDING 624 624
FT ACT_SITE 794 794
FT MOD_RES 825 825
FT CARBOHYD 94 94
FT CARBOHYD 130 130
FT CARBOHYD 145 145
FT CARBOHYD 284 284
FT CARBOHYD 294 294
FT CARBOHYD 301 301
FT CARBOHYD 321 321
FT CARBOHYD 353 353
FT CARBOHYD 368 368
FT CARBOHYD 401 401
FT CARBOHYD 464 464
FT CARBOHYD 487 487
SQ SEQUENCE 978 AA; 109722 MW; CA4D663F98205CA9 CRC64;

Query Match 24.6%; Score 1295.5; DB 1; Length 978;
Best Local Similarity 33.3%; Pred. No. 4.2e-81;
Matches 335; Conservative 169; Mismatches 339; Indels 163; Gaps 35;

Qy 47 VKSSSYVPWSESPDLGC-ALRPQSGTYEAAAEVDVVSASITLOLVDPAGNISCLW 105
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 20 VQTGSGQPSV--SPGELSPLSIHPAKS-----ELIVSGDEIRLLCTDGFVK--W 66
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 106 VFKHSLNCQPHFDLQNRGVSVMLKMTTQAGEYLLFTQSEAT---NYTILFTVSI RN 162
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 67 TFE-----ILQLSEKTNPEWIT-EKAEATNTGNTCTNKGGLSS 105
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 163 TL-----LYTLRRPYFRKMENQDALVCISVPEPIVEWVLCDSQGSKKEESP AV 213
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 106 SIYVFVRDPEKFLIDLPLVYKKEEN-DTLVRCPLTDPE-VTNYSLTGCEGKPLKDLTFV 163
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 214 -----VKKEKVLHFLFGTDIRC--CARNELGRE-CTRLFTID----LNQTPQTLP 258
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 164 ADPKAGITIRNVKREYHRL-----CLHGSANOKGKMSLKFTLVKRAIKAVPVVSVS 217
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 259 Q--LFLKVGELPWIRCKAVVNHGFLGTWELENKALEEGNYFEMSTYSTNRTWIRILFAF 316
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 218 KTSYLLREGEFAVTCILIKDVSSVDSVMWIKENSSQSKAQTKKNSHQGDFSYLRQERLT 277
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 317 VSSVARNDTGYTTCSSSKHPQSALVT---IVGKGFINA-----TNSSDEYDIQY 364
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 278 ISSARVNDSGVFCYANNFTFGSANVTTLEVDVKGINIFPMNNTTVFVNDGENVDL--- 334
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 365 EBFCSFVRFKAYPQ-----IRCTWTSRKSFPCEQKGLDNGYSISKPCNHKH----- 411
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 335 -----VVEYEAVPKPEHRQWIVNRTSTDKWDYPKSE-----NESNIRVYNELHLTRLK 384
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Oy 412 --OPGEYIFHAENDDAOFTMTNIRRRKPOVLAESA--SOASCSDGYPPLPSWTWKC 467
Db 385 GTEGGTYTHFVNSDVSNSSTFNNVNTPEILTHDRLVNGLMQCVAAGFPETIDWYFC 444
Oy 468 SDKSPNCTBEI--TEGVNKRANKRVGQWSSSTLMSAIGFLVKCCAYNSLGTSCRT 526
Db 445 PTEORCSPVGPVDVQIONSSVSPKLVVSTIDDSFTFKNGTVECRAYNDVGKSSAS 504
Oy 527 ILLNSPGPPFFIQDN-----ISFYATIGVCLLFFVLTLLICHYKQKQFY 572
Db 505 F-----NFAFGKNKEQIHAHTLFTPLLGIFVIAAGLMCIFMILT-----YKLLQKPMY 554
Oy 573 ESQLOMV-QVTSSDNEYFYVDFREYEDLWKEFFRENLEFGKVLGSGAFGKMMATAYG 631
Db 555 EYQWKVVEING---NNYVYIIDPTQLPFDHKWEFFPRNLSFGKTLGAGAFGKVEATAYG 611
Oy 632 ISKTGVSIOAVKMLKEKADSEREALMSLQMTOLGSHENIVNLLGACTLSGPIYLIIF 691
Db 612 LIKSDAAMTVAVKMLKPSAHLTEREALMSLKVLSVLGNHMIIVNLLGACTIGPFLVIT 671
Oy 692 EYCCYGDLLNLYRSRREKPHRTWTE-----IPKEHNFYSYPTFQSHPNSSM---PGSRE 742
Db 672 EYCCYGDLLNLFRRKEDSFICKQEDHAELVALYKLLHSSKSSCNDSTNEYMDKPGVSY 731
Oy 743 VOIHPSDQISGLHNSFHSBEDIEYENOKRLEEBEDNLVTFEDLLCTAYQVAKGMFL 802
Db 732 VVPTKAAADRKRSGARIGSY-----IERDVTAPIMEDDEL-ALDLEDLSFSYQVAKGMFL 785
Oy 803 EFKSCVHRDLAARNVLTHGKVKVVICDFGLARDIMSDSNVYVGNARLPVKMAPESLFE 862
Db 786 ASKNCHIRDLAARNILLTHGRITKICDFGLARDIKNDSNVYVGNARLPVKMAPESIEN 845
Oy 863 GYITIKSDVWSYGLLWEIFSLGVPYGPVDPANFYKLIQNGFKMDQPFYATEEYIIM 922
Db 846 CVYTESDWSYGIFLWELFSLGSSPYGMPVDSKYKMKEGFRMLSPHAPENYDIM 905
Oy 923 QSCWAFDSKRSFPLNLSFLGCOLADAEAMVQNVYDGRVSEC-PH 967
Db 906 KTCWDADPLKPTFKQIVQLIEKQISESTNHIYSN-----LANGSPH 947

RESULT 4
KIT_BOVIN
ID KIT_BOVIN STANDARD; PRT; 977 AA.
AC P43481.
DT 01-NOV-1995 (Rel. 32, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)
DE (Proto-oncogene tyrosine-protein kinase kit) (c-kit).
GN KIT.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelestomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]_
RP SEQUENCE FROM N.A.
RX MEDLINE=94215924; PubMed=7512939;
RA Kubota T., Hikono H., Sasaki E., Sakurai M.;
RT "Sequence of a bovine c-kit proto-oncogene cDNA.";
RL Gene 141:305-306(1994).
CC -!- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
CC 3-KINASE (PI3K).
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D16680; BAA04084.1; -.
CC HSSP: P11362; 1FGK.
CC InterPro: IPR007110; IG-like.
CC InterPro: IPR003599; IG.
CC InterPro: IPR003006; IG_MHC.
CC InterPro: IPR000719; Prot_kinase.
CC InterPro: IPR001824; RTKinaseIII.
CC InterPro: IPR001245; Tyr_kinase.
CC Pfam: PF00047; ig; 1.
CC Pfam: PF00069; pkinase; 1.
CC ProDom: PD000001; Prot_kinase; 2.
CC SMART: SM00409; IG; 3.
CC SMART: SM00219; TyRK; 1.
CC PROSITE: PS00835; IG LIKE; 2.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
CC PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
CC Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
CC Transferrase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
CC Immunoglobulin domain.
CC SIGNAL 1 22
CC CHAIN 23 977
CC DOMAIN 23 521
CC TRANSMEM 522 544
CC DOMAIN 545 977
CC DOMAIN 590 938
CC NP_BIND 596 604
CC BINDING 624 624
CC ACT_SITE 793 793
CC MOD_RES 824 824
CC CARBOHYD 94 94
CC CARBOHYD 130 130
CC CARBOHYD 145 145
CC CARBOHYD 284 284
CC CARBOHYD 294 294
CC CARBOHYD 301 301
CC CARBOHYD 321 321
CC CARBOHYD 353 353
CC CARBOHYD 368 368
CC CARBOHYD 401 401
CC CARBOHYD 464 464
CC CARBOHYD 487 487
CC SEQUENCE 977 AA; 109685 MW; 4B2719050883B7EF CRC64;
Query Match 24.4%; Score 1286; DB 1; Length 977;
Best Local Similarity 33.2%; Pred. No. 1.9e-80;
Matches 332; Conservative 172; Mismatches 344; Indels 152; Gaps 36;
Oy 47 VKKSSYPWVSESPDLGC-ALRPOSSGTVYAAAVEVDVSAISITQVLVDAGNLSCLW 105
Db 20 VQTGSSQPSV--SPGELSIPSHPAKS-----ELIVSGDEIRLLCTDGFVK--W 66
Oy 106 VPKHSSLNCQPHFDLQNRGVSMVLKMTQAGEVLLFIOSEAT---NYTLFTVTSIRN 162
Db 67 TFE-----ILQSEKTNPEWIT-EKAEATNTGNTCTNKGGLSS 105
Oy 163 TL-----LYTLRRPYFRKMENQDALVCISESVPEPIVEWVLCDSCSKESPAV 213
Db 106 SIYVVRDPEKFLDLDPLYGKEEN-DTLVRCELTDPV-VTNYSLTGCSEKPLPKDLTFV 163
Oy 214 -----VKKEKVLHFLGTDIRC--CARNELGRE-CTRLFTID-----LNTPQTLTP 258
Db 164 ADPKAGITIRNVKREYHRL-----CLHCSANQRGKSMLSKFKTLKVRRAIKAVPVSVS 217

```

```
QY 259 Q--LFLKVGEPWIRCKAVVNHGFLGTWELENKALEEGNYFEMSYNRTWIRLILFAP 316
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 218 KTSYLLREGEFAVTCIKDVSSVSDMWIKENSQQTAKQTKNSHQGDFSYLRQERLT 277
QY 317 VSSVARNDTGYTCSSSKHPSQSALVT---IVKGFEINA-----TNSSEDIYDQY 364
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 278 ISSARVNDGVMFCYANFTGSAVNTTLEVDKGFNIPFMMNTTVFVNDGENVDL--- 334
QY 365 BEFCFVRKAYPO-IRCTWTSTRKS-----FPCQKGLDNGYSGISFCNKH- 411
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 335 -----VVEEAPKPVHROWIYNRTSTDKWDYPKSE-----NESIRVYNELHLTRLK 384
QY 412 --QPGEYIIPHAENDDAOFTKMTFLNIRRPQVLAESA--SQASCPSDGYPLPSWTWKKC 467
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 385 GTEGGTYTFHVSNSDVNSVTFVYVNTKPEILTHDLVNGMLQCVAAGFPEPTIDWYFC 444
QY 468 SKSPNCTREI-TEGVNRKANRKFVGQWSSSTLNMSKALGFLYKCCAYNSLGTSCET 526
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 445 PGTQRCSPVPGVDVQIQNSSVSPFGKLVVYSTIDDTFKHNGTVECRAYNDVGKSSAS 504
QY 527 ILLNSPGPPP-----FIQDNISVATIGVCLLFIIVLTLCHIKYKQFYESQLOM 578
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 505 FNFAFKGNSKEQIHAHTLFTPLLIGFVIAAGLMCIFWILT-----YKLOKPMYEVQWKV 560
QY 579 V-QVTGSSDNEYFYVDFREYEDLKWEPRENLEFGKVLGSGAGFGKVMNATAYGISKTGV 637
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 561 VEEING--NNVYVIDPTQLPYDHKWEPFNRNLSFGKTLGAGAFGKVEATAYGLIKSDA 617
QY 638 STQAVAKMLKEKADSSEREAELKMTQOLGSHENIVNLLGACTLSGPIYLIFFECVYG 697
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 618 AMTVAKMLKPSAHLTEREALMSKLVKLSYLGHNHMINVLLGACTTGGPTLVITECCYG 677
QY 698 DLLNLYLSRKREKPHRTWTE-----IFKEHNFYSYFTFQSHPNSSM---PGSREVOIHDP 748
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 678 DLLNFRRRKDSFICQEDHAEVALYKLLHKSSEKNDSTNEYMDKRGVSIV-VPTK 736
QY 749 SDQISGLHNSHSEDEIYENOKRLEEBDLNVLTFEDLLCFAYQVAKGMEFLFKSCV 808
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 737 ADKRRSARIGSY-----IERDVTPAIMWEDDEL-ALDLEDLLSFSYQVAKGMFLASKNCI 790
QY 809 HRDLARNVLVTHGVKVKICDFGLARDIMSDSNVYVGRNARLPVKWAPESLFEGIYTIK 868
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 791 HRDLARNILLTHGRITKICDFGLARDIKNDSNVYVGNARLPVKWAPESIFNCVYTFE 850
QY 869 SDVWSYGILLWEIYFSLGVNPGIPVDANFYKLIQNGKMDQPFYATEIYIIMQSCWAF 928
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 851 SDVWSYGIFLWEIYFSLGVNPGIPVDANFYKLIQNGKMDQPFYATEIYIIMQSCWAF 910
QY 929 DSRKRPSFPLNLTSLFLGLADAEAMYNQVNDGRVSEC-PH 967
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 911 DPLKRTFTKQIVQLIEKQISESTNHIYSN---LANCSPH 946

RESULT 5
KIT MOUSE
ID KIT MOUSE STANDARD; PRT; 975 AA.
AC P05532; Q61415; Q61416; Q61417;
AC 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)
DE (Proto-oncogene tyrosine-protein kinase Kit) (c-kit).
GN KIT OR SL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=BAIB/c;
RX MEDLINE=88296403; PubMed=2456920;
RA Qiu F., Ray P., Brown K., Barker P.E., Jhanwar S., Ruddle F.H.,
```

```
RA Besmer P.;
RT "Primary structure of c-kit: relationship with the CSF-1/PDGF
RT receptor kinase family -- oncogenic activation of v-kit involves
RT deletion of extracellular domain and C terminus.";
RL EMBO J. 7:1003-1011(1988).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM TRUNCATED).
RC STRAIN=ICR;
RX MEDLINE=92331813; PubMed=1378413;
RA Rossi P., Marziani G., Albanesi C., Charlesworth A., Gerenia R.,
RA Sorrentino V.;
RT "A novel c-kit transcript, potentially encoding a truncated receptor,
RT originates within a kit gene intron in mouse spermatids.";
RL Dev. Biol. 152:203-207(1992).
RN [3]
RP LIGAND.
RX MEDLINE=91006023; PubMed=1698611;
RA Tan J.C., Buck J., Levi E., Besmer P.;
RT "Candidate ligand for the c-kit transmembrane kinase receptor: KL, a
RT fibroblast derived growth factor stimulates mast cells and erythroid
RT progenitors.";
RL EMBO J. 9:3287-3294(1990).
RN [4]
RP VARIANT W42 ASN-790.
RX MEDLINE=90100577; PubMed=1688471;
RA Tan J.C., Nocka K., Ray P., Traktman P., Besmer P.;
RT "The dominant W42 spotting phenotype results from a missense mutation
RT in the c-kit receptor kinase.";
RL Science 247:209-212(1990).
RN [5]
RP VARIANTS W37 LYS-582; WV MET-660 AND W41 MET-831.
RX MEDLINE=90269214; PubMed=1693331;
RA Nocka K., Tan J.C., Chiu E., Chu T.Y., Ray P., Traktman P.,
RA Besmer P.;
RT "Molecular bases of dominant negative and loss of function mutations
RT at the murine c-kit/white spotting locus: W37, W41 and W.";
RL EMBO J. 9:1805-1813(1990).
CC -I- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
CC 3-KINASE (PI3K).
CC -I- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -I- SUBCELLULAR LOCATION: Type I membrane protein.
CC -I- DISEASE: WHITE-SPOTTING VARIANTS INDUCES SEVERE EFFECTS ON
CC PIGMENTATION, GAMETOGENESIS AND HEMATOPOIESIS. MICE HOMOZYGOUS
CC FOR W42 DIE PERINATALLY OF MACROCYTIC ANEMIA.
CC -I- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -I- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Y00864; CAA68772.1; -.
CC EMBL; X65997; CAA46798.1; -.
CC EMBL; X65998; CAA46799.1; ALT_SEQ.
CC EMBL; X65998; CAA46800.1; -.
CC FIR; A44876; A44876.
CC FIR; S00474; TVMSKT.
CC HSSP; P13362; IFGK.
CC MGD; MGI:96677; Kit.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003598; Ig_c2.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR000719; Prot kinase.
CC InterPro; IPR001824; RTKinaseIII.
```

DR InterPro; IPR001245; Tyr\_kinase.  
 DR Pfam; PF00047; ig; 1.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR ProDom; PD000001; Prot\_kinase; 2.  
 DR SMART; SM00408; IGC2; 1.  
 DR SMART; SM00219; TyrcK; 1.  
 DR PROSITE; PS00835; IG LIKE; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
 KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;  
 KW Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;  
 KW Immunoglobulin domain; Disease mutation.  
 FT SIGNAL 1 22  
 FT CHAIN 23 975 MAST/STEM CELL GROWTH FACTOR RECEPTOR.  
 FT DOMAIN 23 519 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 520 542 POTENTIAL.  
 FT DOMAIN 543 975 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 588 935 PROTEIN KINASE.  
 FT NP\_BIND 594 602 ATP (BY SIMILARITY).  
 FT BINDING 622 622 ATP (BY SIMILARITY).  
 FT ACT\_SITE 790 790 BY SIMILARITY.  
 FT MOD\_RES 821 821 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT CARBOHYD 146 146 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 296 296 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 303 303 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 323 323 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 355 355 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 370 370 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 466 466 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 489 489 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT VARIANT 1 12 MRGARGAWDLIC -> MAVAVFPPLPQQ (IN TRUNCATED FORM).  
 FT VARIANT 13 785 MISSING (IN TRUNCATED FORM).  
 FT VARIANT 582 582 E -> K (IN W37 SPOTTING).  
 FT VARIANT 660 660 T -> M (IN W4 SPOTTING).  
 FT VARIANT 790 790 D -> N (IN W42 SPOTTING).  
 FT VARIANT 831 831 V -> M (IN W41 SPOTTING).  
 SQ SEQUENCE 975 AA; 109001 MW; BABSCAD9AF9CD2A CRC64;  
 Query Match 24.0%; Score 1266; DB 1; Length 975;  
 Best Local Similarity 33.2%; Pred. No. 4.4e-79;  
 Matches 336; Conservative 173; Mismatches 330; Indels 172; Gaps 41;  
 QY 35 CVLIHNKNDSSVSGKSSYPWVSESPDLGALRQSSSGTYEAAAVEVDVSASITLQVL 94  
 DB 12 CVLLVLLRGQTATSPASPGEPSP-----SIHPAQ-----ELIVEAGDTL--- 54  
 QY 95 VDAPGNISCL-----WFKHSSLNQPHFDLQNRGVVSVILKMTQTQAGEYLLFTQSE 148  
 DB 55 -----SLTCDIDPDVFRWTFK-----TYFN-----EMVENKKNWIIQ-EKAE 89  
 QY 149 ATNYTLFTVSIKNTL-----LYTLRPVFRKMNODALVCISSEYFEPIVE 195  
 DB 90 ATR-TGTYTCNSNGLSYIYFVRDPAKLVGLPLFGK-EDSDALVRCPLTDPQ-VSN 146  
 QY 196 WVLCDSDGECSEKESPV-----VKKEKVLHFLFGTDIRCCARNELGRECTRL--- 244  
 DB 147 YSLIEDGKSLPTDLTFVNPKNPAGITIKNVKRAVRLC-----VRCAQ-----RDGTWLHSD 199  
 QY 245 -FTIDLNQ-----PQTLPLQ-----FLKVGEPILWIECKAVHNVHGFGLTW-----E 286  
 DB 200 KFTLKVREAIKAPVSVSPETSHLLKGGDTFTVVCTIKDVSTSVSNMVKMNPQPHIAQ 259  
 QY 287 LENKALEEGNYFEMSTSTNRTMIRILFAPVSSVARNDTGYITCCKSPQSALVT--- 343  
 DB 260 VKNSWHRGDF-----NYEROETLT-----ISSARVDDSGVFMCYANNFTGSAVNTTLK 309  
 QY 344 IVKGGFINATN--SSEDEIDQYEFECFVSRKAYPQ-IRCTWTFSRKSPCEQKGLD--- 398  
 DB 310 VVEKGFINISPVKNTTFTVDGENVDLVVEYEAIPKPEHQWQIYNNRT--SANKGKYVK 367

QY 399 -NGYSISKFCNKH-----QPGEYIFHAENDDAQFTKMTLIRKQPQVLA--EASAS 448  
 DB 368 SDNKSNIYVYNQLRLTLKGTGGTYTFLVSNSDASASVTENVYNTPEILTYDRLNG 427  
 QY 449 QASCFSQGYPLPSWTKKCDKSPNCTEETITE-GVMNRKANRKFVGQWSSSTLMSAI 507  
 DB 428 MLQCAVAGFPEPTIDWYFCTGAEQRCTTPSPVDVQNVSVSPFGKLWQSSIDSSVFR 487  
 QY 508 KGFLVKCCAYNSLGTSCETILLNSPGPPF---IQDN-----ISFVATIGVCLLFIW 557  
 DB 488 HNGTVECKASNDVGKS--SAFFN---FAFKEQIQAHTLFTPLLIGFVVAAGAMGIIWV 541  
 QY 558 LTLILCHYKKQPRYEQSOLQWV-OVTGSSDNEYFYVDPREYEDLKWEPFRENLEFGKVL 616  
 DB 542 LT-----YKLOKPMYEVQWKEEING---NNYVIIDTQPLPDHUKWEPFNRNLSFGKTL 594  
 QY 617 GSGAFGKVMNATAYGISKTGVSIOVAVKMLKESEREAALMSLKMOTQSGHENIVN 676  
 DB 595 GAGAFGKVEATAYGLIKSDAAMTAVKMLKPSAHLTEREALMSLKVLSYLGHNHIVN 654  
 QY 677 LLGACTLSGFIYLIIFYCCYCGDGLLNLRSKREK-----HRTWTEIFKEHNFSPYTFQ 730  
 DB 655 LLGACTVGGPTLVITYCCYCGDGLLNLRRKRDSEIFSKQEQAAALYKLLHSTEPSCD 714  
 QY 731 SHPNSSM---PGSREVOIHPDSDOI SGLHNGSHSEDEIEYENOKRLEEDDLNVLTFED 787  
 DB 715 S-SNEYMDMKPGSYV-VPTKIDRRSARDISY-----IERDVTPAIMDEDEL-ALDDDD 766  
 QY 788 LLFCAYQVAKMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIMSDSNVVRGN 847  
 DB 767 LLSFSYQVAKMAFLASKNCIHRDLAARNILLTHGRITKICDFGLARDIRDSNVVVKGN 826  
 QY 848 ARLPVKWAPESIFEGIYTKSDVMSYIGILLWEIFSLGVNPPYGPVDPVANYKILQNGFK 907  
 DB 827 ARLPVKWAPESIFSCVYTFESDVMYSGIFLWELFSLGSSPYGMPGVDSEKFKYKIKEGR 886  
 QY 908 MDOPPYATEIYIIMQSCWAFDSKRPSPFNLTSLFGLQCLADAEAMQNV 958  
 DB 887 MVSPEHAPAEYMDVMKTCWDADPLKRPTRFKQVQLIEKQISDSKTHIYSNL 937  
 RESULT 6  
 KIT\_HUMAN  
 ID KIT\_HUMAN STANDARD; PRT; 976 AA.  
 AC P10721;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-JUL-1989 (Rel. 11, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)  
 DE (Proto-oncogene tyrosine-protein kinase Kit) (c-kit) (CD117 antigen).  
 GN KIT.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Fetal brain, and Term placenta;  
 RX MEDLINE=88111521; PubMed=2448137;  
 RA Yarden Y., Kuang W.-J., Yang-Feng T., Coussens L., Munitz S.,  
 RA Dull T.J., Chen E., Schlessinger J., Francke U., Ullrich A.;  
 RT "Human proto-oncogene c-kit: a new cell surface receptor tyrosine  
 RT kinase for an unidentified ligand.";  
 RL EMBO J. 6:3341-3351(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93064697; PubMed=1279499;  
 RA Giebel L.B., Strunk K.M., Holmes S.A., Spritz R.A.;  
 RT "Organization and nucleotide sequence of the human KIT (mast/stem  
 RT cell growth factor receptor) proto-oncogene.";  
 RL Oncogene 7:2207-2217(1992).  
 RN [3]  
 RP SEQUENCE FROM N.A.

RX MEDLINE=97179223; PubMed=9027509;  
 RA Andre C., Hanpe A., Lacharme P., Martin E., Wang X.P., Manus V.,  
 RA Hu W.X., Galibert F.;  
 RT "Sequence analysis of two genomic regions containing the KIT and the  
 RT FMS receptor tyrosine kinase genes.";  
 RL Genomics 39:216-226(1997).  
 RN [4]  
 RP VARIANT LYS-583.  
 RX MEDLINE=92291284; PubMed=1376329;  
 RA Fleischman R.A.;  
 RT "Human piebald trait resulting from a dominant negative mutant allele  
 RT of the c-kit membrane receptor gene.";  
 RL J. Clin. Invest. 89:1713-1717(1992).  
 RN [5]  
 RP VARIANT LEU-584.  
 RX MEDLINE=92133600; PubMed=1370874;  
 RA Spritz R.A., Giebel L.B., Holmes S.A.;  
 RT "Dominant negative and loss of function mutations of the c-kit  
 RT (mast/stem cell growth factor receptor) proto-oncogene in human  
 RT piebaldism.";  
 RL Am. J. Hum. Genet. 50:261-269 (1992).  
 RN [6]  
 RP VARIANT ARG-664.  
 RX MEDLINE=92020918; PubMed=1717985;  
 RA Giebel L.B., Spritz R.A.;  
 RT "Mutation of the KIT (mast/stem cell growth factor receptor)  
 RT proto-oncogene in human piebaldism.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:8696-8699(1991).  
 RN [7]  
 RP VARIANT VAL-816.  
 RX MEDLINE=94013473; PubMed=7691885;  
 RA Furitsu T., Tsujimura T., Tono T., Ikeda H., Kitayama H.,  
 RA Koshimizu U., Sugahara H., Butterfield J.H., Ashman L.K.,  
 RA Kanayama Y., Matsuzawa Y., Kitamura Y., Kanakura Y.;  
 RT "Identification of mutations in the coding sequence of the proto-  
 RT oncogene c-kit in a human mast cell leukemia cell line causing  
 RT ligand-independent activation of c-kit product.";  
 RL J. Clin. Invest. 92:1736-1744 (1993).  
 RN [8]  
 RP VARIANTS PIEBALDISM GLY-791 AND VAL-812.  
 RX MEDLINE=93322624; PubMed=7687267;  
 RA Spritz R.A., Holmes S.A., Itin P., Kuester W.;  
 RT "Novel mutations of the KIT (mast/stem cell growth factor receptor)  
 RT proto-oncogene in human piebaldism.";  
 RL J. Invest. Dermatol. 101:22-25(1993).  
 RN [9]  
 RP VARIANT PIEBALDISM 893-GLU-PRO-896 DEL.  
 RX MEDLINE=96287384; PubMed=9680409;  
 RA Riva P., Milani N., Gandolfi P., Larizza L.;  
 RT "A 12-bp deletion (7818del12) in the c-kit protooncogene in a large  
 RT Italian kindred with piebaldism.";  
 RL Hum. Mutat. 6:343-345(1995).  
 RN [10]  
 RP VARIANT GIST VAL-559 DEL.  
 RX MEDLINE=98361155; PubMed=9697690;  
 RA Nishida T., Hirota S., Taniguchi M., Hashimoto K., Isozaki K.,  
 RA Nakamura H., Kanakura Y., Tanaka T., Takabayashi A., Matsuda H.,  
 RA Kitamura Y.;  
 RT "Familial gastrointestinal stromal tumours with germline mutation of  
 RT the KIT gene.";  
 RL Nat. Genet. 19:323-324 (1998).  
 CC -1- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL  
 CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.  
 CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND  
 CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL  
 CC 3-KINASE (PI3K).  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- DISEASE: DEFECTS IN KIT ARE THE CAUSE OF A DEVELOPMENTAL  
 CC ABNORMALITY KNOWN AS PIEBALDISM. AN AUTOSOMAL DOMINANT GENETIC  
 CC DISORDER OF PIGMENTATION CHARACTERIZED BY CONGENITAL PATCHES OF  
 CC WHITE SKIN AND HAIR THAT LACK MELANOCYTES.

CC -1- DISEASE: DEFECTS IN KIT ARE THE CAUSE OF A GASTROINTESTINAL  
 CC STROMAL TUMOR (GIST).  
 CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
 CC PROTEIN KINASES.  
 CC -1- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.  
 CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD117 entry;  
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd117.htm".  
 CC -1- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;  
 CC WWW="http://www.infobio.gen.fr/services/chromancer/Genes/KITID127.html".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; X06182; CAA29548.1; -;  
 CC EMBL; X69301; CAA49159.1; JOINED.  
 CC EMBL; X69302; CAA49159.1; JOINED.  
 CC EMBL; X69303; CAA49159.1; JOINED.  
 CC EMBL; X69304; CAA49159.1; JOINED.  
 CC EMBL; X69305; CAA49159.1; JOINED.  
 CC EMBL; X69306; CAA49159.1; JOINED.  
 CC EMBL; X69307; CAA49159.1; JOINED.  
 CC EMBL; X69308; CAA49159.1; JOINED.  
 CC EMBL; X69309; CAA49159.1; JOINED.  
 CC EMBL; X69310; CAA49159.1; JOINED.  
 CC EMBL; X69311; CAA49159.1; JOINED.  
 CC EMBL; X69312; CAA49159.1; JOINED.  
 CC EMBL; X69313; CAA49159.1; JOINED.  
 CC EMBL; X69314; CAA49159.1; JOINED.  
 CC EMBL; X69315; CAA49159.1; JOINED.  
 CC EMBL; X69316; CAA49159.1; JOINED.  
 CC EMBL; U63834; AAC50968.1; -;  
 CC PIR; S01426; TVHUKT.  
 CC HSP; P11362; IFGK.  
 CC Genew; HGNC:6342; KIT.  
 CC MIM; 164920; -;  
 CC MIM; 172800; -;  
 CC MIM; 606764; -;  
 CC GO; GO:0004716; F:receptor signaling protein tyrosine kinase . . .; TAS.  
 CC GO; GO:0007165; P:signal transduction; TAS.  
 CC InterPro; IPR007110; Ig-like.  
 CC InterPro; IPR003598; Ig C2.  
 CC InterPro; IPR003006; Ig MHC.  
 CC InterPro; IPR000719; Prot kinase.  
 CC InterPro; IPR001824; RTKinaseIII.  
 CC InterPro; IPR001245; Tyr\_pkinase.  
 CC Pfam; PF00047; Ig; 1.  
 CC Pfam; PF00069; pkinase; 1.  
 CC ProDom; PD000001; Prot kinase; 2.  
 CC SMART; SM00408; IGC2; 1.  
 CC SMART; SM00219; Tyrc; 1.  
 CC PROSITE; PS00835; IG LIKE; 1.  
 CC PROSITE; PS01017; PROTEIN KINASE ATP; 1.  
 CC PROSITE; PS00011; PROTEIN KINASE DOM; 1.  
 CC PROSITE; PS00109; PROTEIN KINASE TYR; 1.  
 CC PROSITE; PS00240; RECEPTOR TYR KIN III; 1.  
 CC Proto-oncogene, Tyrosine-protein kinase; Receptor; Transmembrane;  
 CC Transferase; Glycoprotein; Phosphorylation; Atp-binding; Signal;  
 CC Immunoglobulin domain; Disease mutation.  
 CC SIGNAL 1 22  
 CC CHAIN 23 976 MAST/STEM CELL GROWTH FACTOR RECEPTOR.  
 CC DOMAIN 23 520 EXTRACELLULAR (POTENTIAL).  
 CC TRANSMEM 521 543 POTENTIAL.  
 CC DOMAIN 544 976 CYTOPLASMIC (POTENTIAL).  
 CC NP BIND 589 937 PROTEIN KINASE.  
 CC BINDING 623 623 ATP (BY SIMILARITY).  
 CC ACT\_SITE 792 792 ATP (BY SIMILARITY).  
 CC MOD\_RES 823 823 BY SIMILARITY.  
 CC PHOSPHORYLATION (AUTO-) (BY SIMILARITY).

```

FT CARBOHYD 130 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 145 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 283 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 320 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 367 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 486 MISSING (IN GIST).
FT VARIANT 559 /FTID=VAR 007965.
FT VARIANT 583 E -> K (IN PIERBALDISM).
FT VARIANT 584 F -> L (IN PIERBALDISM).
FT VARIANT 664 /FTID=VAR 004105.
FT VARIANT 791 G -> R (IN PIERBALDISM).
FT VARIANT 812 R -> G (IN PIERBALDISM).
FT VARIANT 816 /FTID=VAR 004107.
FT VARIANT 816 G -> V (IN PIERBALDISM).
FT VARIANT 816 /FTID=VAR 004108.
FT VARIANT 816 D -> V (IN MAST CELL LEUKEMIA;
FT CONSTITUTIVELY ACTIVATED).

Query Match 23.7%; Score 1251; DB 1; Length 976;
Best Local Similarity 32.7%; Pred. No. 4.7e-78;
Matches 324; Conservative 166; Mismatches 352; Indels 148; Gaps 34;

QY 47 VKSSSYPMVSPEDLGCALRQSGTGVYEAADVVSASITLQVLVDAPCNISCLVW 106
DB 20 VOTGSSQSVSPG-BPSPPSIHPGKSDL-----VRVGDRIKLLCTDGFVK-WT 67
QY 107 FKHSSLNQPHFDLQNRGVVSVILKMTOTAGEYLLFIQSEATN---YTLFTVSIRNT 163
DB 68 FE-----ILDETENKQNEWT-EKAEATNGKYCTCNKGLSNS 106
QY 164 L-----LVTLRPYPRKVENQDALVCISBPVPEVWVLCDSQGESKBE----- 209
DB 107 IYVVRDPAKFLVDRSLYK-EDNDTLVRCPLTDP-E-VTNYSLKCGCKPLPKOLRFP 164
QY 210 ---SPAUVKKEKVLHELFGTDIIRCARNELGRECTRLFTIDL---NOTQTLLPQ--L 260
DB 165 DPKAGIMKSVKRAHRLC---LHCSVDQEGKSVLSEKFIKVRPAKFAVPVSVSKASY 221
QY 261 FLKVGPELWIRKCAVNVNHGFLTWELN---KALBEGNVFEMSTYSTNRTMIRILFAFV 317
DB 222 LLREGEEFTVCTIKDVSSSVYTWKRENSQTKLOEKYNSWHGDPNYERQAT-----LTI 277
QY 318 SSVARENDGYTTCSSSKHPSQSALVT---IVGKGFINA-----TNSSEDEYIDQYE 365
DB 278 SSARVNDSGVFCYANNTFGSANVTTLVVDKGFNIPFIMINTTVFVNDGENVDL--- 333
QY 366 ECFVSURKAYPQ-IRCTWTFBRKSP--PCBQKGLDNGYSISKFNKHI-----OPGE 415
DB 334 ---IVEYEAPKPEHQOQIYNNRTFTDKWEDYPKSENEISNRYVSELHLRLKGTGEGT 389
QY 416 YIFHAENDDAQFTKFTLIRKPKQVLA--EASASQASCFSQCYPLPSTWTKKCDKSPN 473
DB 390 YTLFVNSNDVNAIAFNVTYNTKPELTYDRLVNGMLQCVAGFPPTIDWFCPQTEOR 449
QY 474 CTEETE-GVMNRKANRKFVGQWSSSTLNMSEAIKGLVLCVCAVNSLGTSCETILLNSP 532
DB 450 CSASVLPVDVQTLNSSGPPFKGLVQSSIDSSAFKNGTVECKAYNDVGT--SAYFN-- 505
QY 533 GFPPFIQDN-----ISYATIGVCLLFIVVLTLICHYKKQPRYESQLQM 578
DB 506 --FAFGKNKEQIHPHTLFTPLLLIGFVIVAGMCIIIVMLT-----YKYLQKPMYEQVKV 559
QY 579 V-QVTGSSDNEYFYVDFREYEDLWKEFPRENLERCKVLGSGAFGKVMNATYGISKTGV 637
DB 560 VEEING---NNVYVDPTQLPDYDHRKWEFPNRLSFGKTLGAGAFGKVEATYGLIKSDA 616

```

```

QY 638 SIQAVKMLKEKADSSEREALMSLKMMTQIOLGSHENIVNLLGACTLSGPIYLIFECYCG 697
DB 617 AMTAVKMLKPSAHLTEREALMSLKVLISYLGHNWIVNLLGACTIGGFLVITEYCCYCG 676
QY 698 DLLNLYLRSKREKPHRTWTE-----IFKEHNPFSFYPTFQSHPNSSM---PGSREVOIHPD 748
DB 677 DLLNLFRRKRDSEFCISKQEDHAEALYKNLLHSKSSSDSTNEYMMDKPGVYV-VPTK 735
QY 749 SDQISGLHNSPHSEDETEYENQKLEBEEDLNVLTFEDLLCFAYOVAKGMEFLEPKSCV 808
DB 736 ADKRSVRIGSY-----IERDVTIPAIMEDEL-ALDLELLSFSYQVAKGMAFLASKNCI 789
QY 809 HDLAARNVLVTHGKVKIKCDFGLARDIMSDSNVYVGRNARLPVKWMAPESLFEGIVYTIK 868
DB 790 HDLAARNVLLTHGRITKICDFGLARDIKDNNVYVKGARUPVKWMAPESFNCVITFE 849
QY 869 SDVMSYGILLWEIFSLGVNYPYFGIPVDANFYKLQNGFKDQDFFVATBEIYIIMQSCWAF 928
DB 850 SDVMSYGIFLWELFSLGSSPYFGMPVDOSKFYKWKIEGFRMLSPHAPAEWYDIMKTCWDA 909
QY 929 DSRKPSFPNLTSLFGCQLADAEEMVQNV 958
DB 910 DPLKPTFKQIVQLIEKQISESTNHIYSNL 939

RESULT 7
KIT CANFA
ID - KIT CANFA STANDARD; PRT; 975 AA.
AC O97799;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)
DE (Proto-oncogene tyrosine-protein kinase Kit) (c-kit).
GN KIT.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99142897; PubMed=9989791;
RA Ma Y., Longley B.J., Wang X., Blount J.L., Langley K., Caughey G.H.;
RT "Clustering of activating mutations in c-KIT's juxtamembrane coding
RT region in canine mast cell neoplasms."
RL J. Invest. Dermatol. 112:165-170(1999).
CC -!- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY..
CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
CC 3-KINASE (PI3K).
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE CSP-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF044249; AAD02327.1; -.
CC HSSP; P11362; 1FGK.
CC InterPro; IPR007110; IG-like.
CC InterPro; IPR003599; IG.
CC InterPro; IPR003006; IG_MHC.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR001824; RTKinaseII.

```



DR	InterPro; IPR001245; Tyr_kinase.	QY	471	SPNCTEEI-TEGVNRKANRKFVQWVSSSTLNMSEAIKGFVLCVCAVNSLGTSCETILL	529
DR	Pfam; PF00047; IG; 2.	Db	450	BQRCVPIGPMVQVQSSLSFSGKLVVQSSIDISAFKHNGTVECTRAVNVGRS--SAFF	507
DR	ProDom; PD000001; Prot_kinase; 2.	QY	530	NSPGPFPIQD-----NISFYATIGVCLLFIIVVLTLLIICHYKQKQFYESQLQVW	579
DR	SMART; SM00409; IG; 3.	Db	508	N----PAFKEQIHPTLPTLLIGFVIAAGWMCIIIMILT---YKYLQKPMYEQWKVV	559
DR	SMART; SM00219; TyrKc; 1.	QY	580	-QVTGSSDNEYFYVDREYBYDLKWEPPREMLFQKVGSGAFGKMMATAYIGISKTGVS	638
DR	PROSITE; PS00107; PROTEIN KINASE ATP; 1.	Db	560	EEING--NNYVYIDPTQLPYDHKWEPPRNRLSFGKTLGAGAFGKVEATAYGLIKSDAA	616
DR	PROSITE; PS00109; PROTEIN KINASE TYR; 1.	QY	639	IQVAVKMLKEVADSSERELMSKMTQLSHENIVNLGACTLSGPIYLIFFCYCYGD	698
DR	PROSITE; PS00240; RECEPTOR TYR KIN III; 1.	Db	617	MTVAVKMLKPSAHLTEREALMSKLSYLGNHNVNLGACTVGGPTLVITECYCYGD	676
KW	Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;	QY	699	LLNVLRSKREKPHRTWTE-----IFKEHNSFYPTFQSHNNSM---PGSREVQIHPDS	749
KW	transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;	Db	677	LLNVLRRKDSFICKQEDHGEVALYKLLSKSSCSDSINEYMDKFGVSIV-VPTKA	735
FT	Immunoglobulin domain.	QY	750	DQISLHGNSFHSEDEIEYENQKLEEDLNVLTFDLFCFAYQVAKMEFLFKSVH	809
FT	1	Db	736	DKRSARISGV-----IERDVTPTAIMEDEL-ALDLEDLSFSYQVAKGMFLASKNCIH	789
FT	2	QY	810	RDLAARNVLVTHGKVKICDFGLARDIMSDSNVYVRGNARLPVKWMAPELFEIYTIKS	869
FT	3	Db	790	RDLAARNILLTHGRITKICDFGLARDIKNDNSVYVYVGNARLPVKWMAPELFEIYTIKS	849
FT	4	QY	870	DWMSGILLWEIFSLGVNPPYGPVDANFYKLIQNGKMDQFPYATEEIIYIMQSCWAPD	929
FT	5	Db	850	DWMSGIFLWELFSLGSSPYGMPVDSKPYKMKEGFRMLSEHAPAEWIDIMKTCWDAD	909
FT	6	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	7	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	8	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	9	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	10	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	11	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	12	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	13	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	14	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	15	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	16	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	17	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	18	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	19	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	20	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	21	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	22	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	23	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	24	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	25	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	26	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	27	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	28	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	29	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	30	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	31	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	32	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	33	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	34	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	35	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	36	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	37	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	38	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	39	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	40	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	41	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	42	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	43	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	44	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	45	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	46	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	47	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	48	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	49	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	50	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	51	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	52	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	53	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	54	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	55	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	56	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	57	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	58	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	59	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	60	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	61	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	62	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	63	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	64	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	65	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	66	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	67	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	68	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	69	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	70	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	71	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	72	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	73	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	74	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	75	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	76	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	77	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	78	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	79	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	80	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	81	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	82	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	83	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	84	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	85	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	86	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	87	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	88	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	89	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	90	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	91	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	92	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	93	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	94	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	95	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	96	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	97	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	98	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	99	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	100	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	101	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	102	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	103	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	104	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	105	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	106	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	107	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	108	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	109	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	110	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	111	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	112	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	113	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	114	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	115	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	116	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	117	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	118	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	119	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	120	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	121	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	122	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	123	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	124	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	125	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	126	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	127	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	128	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	129	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	130	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	131	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	132	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	133	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	134	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	135	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	136	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	137	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	138	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	139	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	140	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	141	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	142	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	143	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	144	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	145	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	146	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	147	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	148	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	149	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	150	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	151	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	152	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	153	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	154	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	155	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	156	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	157	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	158	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	159	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	160	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	161	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	162	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	163	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	164	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	165	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	166	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	167	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	168	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	169	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	170	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	171	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	172	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	173	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	174	QY	930	SRKPSFENLTSFLGCOLADAEAEAMYQNV	958
FT	175	Db	910	PLKRTFQIVQLIEKQISDSTNNHIYSNL	938
FT	176	QY			



tyrosine-substrate interaction site.";  
 CC OncoGene 10:369-379(1995).  
 CC -!- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL  
 CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.  
 CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND  
 CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL  
 CC 3-KINASE (PI3K).  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
 CC PROTEIN KINASES.  
 CC -!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.  
 CC  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announcement/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; S76596; AAB3207.1; -.  
 CC HSPSP; P11362; 1FGK.  
 CC InterPro; IPR007110; Ig-like.  
 CC InterPro; IPR003599; Ig.  
 CC InterPro; IPR003006; Ig MHC.  
 CC InterPro; IPR000719; Prot\_kinase.  
 CC InterPro; IPR001824; RTKinaseII.  
 CC InterPro; IPR001245; Tyr\_kinase.  
 CC Pfam; PF00047; Ig; 1.  
 CC Pfam; PF00069; pkinase; 1.  
 CC ProDom; PD000001; Prot\_kinase; 2.  
 CC SMART; SM00409; IG; 3.  
 CC SMART; SM00219; Tyrc; 1.  
 CC PROSITE; PS50835; IG LIKE; 2.  
 CC PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 CC PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 CC PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 CC PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
 CC ProCo-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;  
 CC Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;  
 CC Immunoglobulin domain.  
 FT SIGNAL 1 22 POTENTIAL.  
 FT CHAIN 23 978 MAST/STEM CELL GROWTH FACTOR RECEPTOR.  
 FT DOMAIN 23 521 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 522 544 POTENTIAL.  
 FT DOMAIN 545 978 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 590 938 PROTEIN KINASE.  
 FT NP\_BIND 596 604 ATP (BY SIMILARITY).  
 FT BINDING 624 624 ATP (BY SIMILARITY).  
 FT ACT\_SITE 793 793 BY SIMILARITY.  
 FT MOD\_RES 824 824 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT CARBOHYD 94 94 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 145 145 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 284 284 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 294 294 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 321 321 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 333 333 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 368 368 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 401 401 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 464 464 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 487 487 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 978 AA; 109449 MW; 6D45472E07440E6B CRC64;

Query Match 23.2%; Score 1224.5; DB 1; Length 978;  
 Best Local Similarity 32.9%; Pred. No. 3.1e-76;  
 Matches 325; Conservative 167; Mismatches 352; Indels 145; Gaps 35;  
 QY 47 VKSSSYPMVSSPEDLGC-ALRPQSGTYEAAAEVDSASITQLVLDAPGNISCLW 105  
 DB 20 YQTGSSQP--SASPGWWSLPSIHPTAS-----ELIVSAGDEIRLLCTDPGFVK--W 66

QY 106 VPKHSSLNCQPHFDLONRGVSVMLKMTQAGEYLLFIQSEAT---NYTIL----- 155  
 DB 67 TFE-----TLGQSEITHNEWIT-EKAEATNTGNTTNGGGLSS 105  
 QY 156 -FTYSIRN-TLLYTLRRPYFRKMNODALVCISESVPEPIVWVLCDSOGESKEESP 213  
 DB 106 SIYVVRDPAPKLPLVDLPYLGK-EDHDTLVRCPLTDPB-VTNYSLRGCGKPLPDLTTF 163  
 QY 214 -----VKKEKVLHFLGTDIRCCA-----RNLGRECTRLFTIDLNOTPTLLPQL- 260  
 DB 164 TDPKAGITIRNVKREYHRLC---LHCSADRKGKSVLSKKFTLKVRAAIRAVPVVSVSKAS 220  
 QY 261 -FLKVGPEPLWRCVAVHNGFGLTWELNKALESGNYFEMSTYSTNRTMIRILFAFVSS 319  
 DB 221 HLLRGESEFVNCCLIKOVSSVDGSMWIKENSPTNAQPSNSWHQDFFNVQRILTSS 280  
 QY 320 VARNDTGYTTCSSSKHPSQSALVT---IVGKGFINA-----TNSSEDEYIDQYEEF 367  
 DB 281 ARVNDGSGVMCVANVTFGSANVTTLVAVAKGFINIFPMWNTTIFVNDGENVDL----- 334  
 QY 368 CFSVRFKAYPQ-----IRCTWTSRKSFPCEQKGLDNGYSISKFCNHHK-----Q 412  
 DB 335 --IVEYEAVPKPEHQRWVVMNRTLTDKWEDYPKS---DNESNI-RYVSELHLTRLKNGE 387  
 QY 413 PGEYIFHAENDDAQFTKMTLNIRKPKQVLAAS--ASQASCFSQGYPLPSPWTKKCSK 470  
 DB 388 GGTTFQVNSDVNSVTLNVVYVTKPILTHESLVSGILOCLVAGFPPEPTVDWVPCGA 447  
 QY 471 SPNCTEEITE-GVMNRKANRKVFGQWSSSTLANMSEAIGFLVKCAVNSLGTSCBTILL 529  
 DB 448 EORCPVPVGPLDVQMNSSVSPGKLVQSSSIDYSAFKINGTVECRASHNVGKT--SAFF 505  
 QY 530 NSPGPFPIQDN-----ISFYATIGVCLLFIIVLTLLIICHKKQKFRYESQ 575  
 DB 506 N----FAFKGSKSQMHPTLFTPLLIGPVIAAGMCIIVMLT---YKYLQKPMYEQ 557  
 QY 576 LQMW-QVTGSSDNEYVDREYEDLKWEPRENLEFGKVLGSGAFGKMMATAYGISK 634  
 DB 558 WKVVEING--NNYVYIDPTOLPYDHKWEFPNRLSFGKTLGAGAFGKVEATAYGLIK 614  
 QY 635 TGVSIQVAVKMLKEKADSSEREAALMSELKMTQGLSHENIVNLGACTLSGPTLILFEYC 694  
 DB 615 SDAAVTAVKMLKPSAHUTEREAALMSELKLVSLYLNHNMVNLGACTVGGPTLVITEYC 674  
 QY 695 CYGDLNLYRSKRERFRTWTETFEKHNFSEFYPT-FQSHPNSSMPGSRV-VQIHPDSQI 752  
 DB 675 CYGDLNLYRKRDSFICSKQEDHAE--VALYKNLLQSKESCNDSNTEYMDMKPGSVY 732  
 QY 753 SGLHNSFHSF---DEIYENQKRLEBEDLVNLTFFEDLLCFAYQVAKMEFLFKSCVH 809  
 DB 733 VPTKADKRSARIGSYIERDVTPTAIMEDEL-ALDLEDLLSFSYQVAKGMFLASKNCIH 791  
 QY 810 RLDAARNVLVTHGKVKLCDFGLARDIMSDNSVYVGNARLVKVMAPESLEGITIKS 869  
 DB 792 RLDAARNILLTHGRITKICDFGLARDIKNDNSVYVGNARLVKVMAPESINFCVTFES 851  
 QY 870 DVMSYGILLWEIFSLGVNPPYGPVDANFYKLIQNGFKMDQFPYATEETIYIMQSCWAFD 929  
 DB 852 DVWSYGFILWELFSLGSSPYFGMPVDSEFYKMKEGFRMLSPHAPAEYDINKTCWDAD 911  
 QY 930 SRKRFPNLTSLFGLCQLADAEANQYV 958  
 DB 912 PLKRPTFKQIVQLIEKQISDSTNHIYSNL 940

RESULT 10  
 KEMS FSVMID  
 ID KEMS FSVMID STANDARD; PRT; 978 AA.  
 AC P00545; Q86597;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)







```

DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001824; TtkinaseIII.
DR InterPro; IPR001245; TtkinaseIII.
DR Pfam; PF00047; Ig_3.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 2.
DR SMART; SM00408; Igc2; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00835; IG LIKE; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
KW Transferrase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
KW Immunoglobulin domain.
FT SIGNAL 1 24
FT CHAIN 25 960
FT DOMAIN 25 960
FT TRANSMEM 506 530
FT DOMAIN 531 960
FT DOMAIN 575 913
FT NP_BIND 581 589
FT BINDING 609 609
FT ACT_SITE 777 777
FT MOD_RES 808 808
FT CARBOHYD 76 76
FT CARBOHYD 135 135
FT CARBOHYD 149 149
FT CARBOHYD 289 289
FT CARBOHYD 286 286
FT CARBOHYD 306 306
FT CARBOHYD 318 318
FT CARBOHYD 338 338
FT CARBOHYD 343 343
FT CARBOHYD 356 356
FT CARBOHYD 453 453
FT CARBOHYD 469 469
SQ SEQUENCE 960 AA; 107311 MW; 0E93850527AB68F6 CRC64;

Query Match 23.0%; Score 1215; DB 1; Length 960;
Best Local Similarity 32.3%; Pred. No. 1.4e-75;
Matches 314; Conservative 165; Mismatches 370; Indels 124; Gaps 34;

QY 66 ALRPOSSGVYEAAYVDDVSAITQLVLDAPGNISCLWVFKHSS-----LNCOPHPD 119
DB 18 SLIPAGSVPHSESSILVNVKGBELRLKCNKEGP-----VTWNQNSDPSAKTRISNEKEWH 73
QY 120 LQNRGV--VSMVILKMTQAGEYLLFIOSEATNTYTLFTVSRNTLLTLRLRPYRKME 177
DB 74 TKNATIRDIGRECKSGKSVNSFYVFKDP-----NVLFV---DSLIV-----GKED 119
QY 178 NQDALVCISVPEP-IVENWLCDGSGECKEESPAVVKKE-----KVLHELFGTDIRCC 232
DB 120 SDILLVC-----PLTDPDLNFTLRKCDKPLKPNMTFIPNPQKGIIITKNVORSFKGCYCL 176
QY 233 AR-NELGRECTRLFTIDLNOTP-OTTLPOL-----FLKVGELPIRCKAVVNHGFGCL 283
DB 177 AKHNGVEKISEHF---LNVRPVHKALPVITLSKVELLKEGEFEVTCIIITDVSUVA 233
QY 284 TWELNKALEEGNYFEMSTYSTNRMTIRILFAFVSSVARNDTGYYTCSSKHP---SQSAL 341
DB 234 SMISYKSAIVTSKSRNLGDYVERK---LTLNIRSVGVNDSGEFTC-QAENPFPGKTNAT 288
QY 342 VTI--VGKGFN---ATNSEDSEIDQYEEFCFVSFKAYPQIR-CTWTFSRKSPFCEQK 395
DB 289 VTLKALAKGFVRLFAFMNTTIDINAGQNGN---LTVVEYAYPKPKKEVVMYMETL---QN 343
QY 396 GLDNGYSISKFNHKK-----OPGEYIFHAENDDAQFTKMTLNTIRRKPOVLA 443
DB 344 SSDHYVKFTVGNNSVTSLSHLTRLKGTGGIYTFVVSNDSSSVTFVNYVTKREIILT 403
QY 444 EASASQ--ASCFSGDGYPLSPWTKKCDKSPNCTEITEGVNVRKANRKFVGQWSSSTL 501

Db 404 LDMGLNDILQCVATGFPAPTYYWFCPTGEQCLDSPMDVKVS-----YTNSSVP 457
QY 502 NMSEAIGKFLYKCCAYNSLGTSCETILLNSPGPPFF-----IQDNISFYA-----TIGV 550
DB 458 SFERILVESTVNASMFKSTGTICCEASSNGDKSSVFFNFAKEQIRTHRTFLTPLLIAGV 517
QY 551 CLLFIVVLTLLICHYKQFQRYESQLQMV-QVTGSSDNEYFYVDREYEDLAKWEPPREN 609
DB 518 AAGLCMIIVMILVIYIQLPKYEQVMKVVEENG---NNYVVIDPTQLPYDHKWEPPNR 574
QY 610 LEFGVILGSGAGFKVMNATAYGISKTGVSIOAVKMLKEKADSSEREAALMSLKMWTQLG 669
DB 575 LSFGLTLAGAGFGKVEATAYGLFKSDAAMTVAVKMLKPSAHLTEREALMSLKVLSYLG 634
QY 670 SHENIVNLGACTLSGPIYLIPEYCCYGDLLNLYRSKREKF-----HRTWTEIFKEHNF 725
DB 635 NHINIVNLGACTIGGPTLVITEYCCYGDLLNLFARKRDSFICPKHEEAAAVYENL-- 692
QY 726 YPTFOSHENS-----SMPGSRVQIHPDSDQISGLHGNSTHSEDEIEYENQKLEEBE 778
DB 693 --LHQAEPTADAVNEYMMDKPGVSYAVPPKADKRPVKSGSYTDQD---VTLSMLEDE 746
QY 779 DLNVLTPEDLILCFAYOVAKGMEFLFKSCVHEDLAARNVLTGKVVKICDCEGLARDIMS 838
DB 747 --LALDVEDLLSFSYQVAKGMSFLASKNCIHRDLAARNILLTHGRITKICDFGLARDIRN 804
QY 839 DSNVYVRGNARLPVKWMAPESLFEGIYITKSDVWVGILLWIFSLGVNPNYPCIPVDANF 898
DB 805 DSNVYVKGARLPVKWMAPESEIFNCVYTFESDVWVGILLWELFSLGSSPYGMPVDSKF 864
QY 899 YKLQNGFMDQPPVATVEIYIIMOSWAFDSKRPSPNLTSLFLGCLADAEAEAMQNV 958
DB 865 YKMIKEGYRMFSPESCPPEMYDIMKSCWDADPLQPTFKQIVQLIEQQLSDNAPRYAN- 923
QY 959 DGRVSECPHTYQN 971
DB 924 ---FSTPPSTQGN 933

RESULT 13
KFMS MOUSE
ID KFMS_MOUSE STANDARD; PRT; 977 AA.
AC P09581; Q9DBH9;
DT 01-MAR-1989 (Rel. 10, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Macrophage colony stimulating factor 1 receptor precursor (CSF-1-R)
DE (EC 2.7.1.112) (fms proto-oncogene) (c-fms).
GN CSF1R OR CSFMR OR FMS;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88217329; PubMed=2966922;
RA Rothwell V.M., Rohrschneider L.R.;
RT "Murine c-fms cDNA: cloning, sequence analysis and retroviral
RT expression.";
RL Oncogene Res. 1:311-324 (1987).
RN [2]
RP REVISIONS.
RA Rothwell V.M.;
RN [3]
RP Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RX MEDLINE=93181280; PubMed=8441691;
RA de Parseval N., Borgeaux D., Giselsbrecht S., Sola B.;
RT "Reassessment of the murine c-fms proto-oncogene sequence.";
RL Nucleic Acids Res. 21:750-750 (1993).
RN [4]
RP SEQUENCE FROM N.A.

```











GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: August 26, 2003, 07:20:28 ; Search time 83.5421 Seconds  
3067.272 Million cell updates/sec

Title: US-09-919-408A-4  
Perfect score: 5274  
Sequence: 1 MPALARDAGTVPLLVVFSAM.....PFSREMDGLLSPOQAQVEDS 993

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues  
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPREMBL 23:\*

- 1: sp\_archea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1707.5	32.4	406	11 Q8BR20	Q8br20 mus musculus
2	1272.5	24.1	976	13 Q8JPR5	Q8jfr5 brachydanio
3	1271.5	24.1	976	13 Q9W755	Q9w755 brachydanio
4	1271	24.1	979	11 Q8C8K9	Q8c8k9 mus musculus
5	1265.5	24.0	974	11 Q8C702	Q8c702 rattus ratt
6	1263.5	24.0	978	11 Q63116	Q63116 rattus norv
7	1262.5	23.9	964	6 Q97744	Q97744 sus scrofa
8	1262	23.9	984	13 Q8AXC6	Q8axc6 fugu rubrip
9	1261.5	23.9	964	6 Q9TQ01	Q9tq01 sus scrofa
10	1261.5	23.9	964	6 Q9TQ00	Q9tq00 sus scrofa
11	1250.5	23.7	979	6 Q8WN23	Q8wn23 canis fami1
12	1247	23.6	972	4 Q99662	Q99662 homo sapien
13	1244	23.6	978	6 Q9XS93	Q9xs93 canis fami1
14	1230.5	23.3	977	13 Q98SU1	Q98sul danio nigro
15	1229	23.3	945	6 Q77589	Q77589 equus cabal
16	1228	23.3	974	13 Q98SU3	Q98su3 danio dangi

17	1226	23.2	954	13 Q91909	Q91909 xenopus lae
18	1223.5	23.2	977	13 Q918N6	Q918n6 brachydanio
19	1221	23.2	948	6 Q9TDT7	Q9tdt7 trichosurus
20	1219	23.1	724	6 Q9MYN0	Q9myn0 bos taurus
21	1214	23.0	977	13 Q98SU2	Q98su2 danio kerri
22	1213	23.0	975	13 P79750	P79750 fugu rubrip
23	1211	23.0	977	13 Q98SU4	Q98su4 danio albol
24	1192.5	22.6	992	13 Q8AXU0	Q8axu0 oncorhynch
25	1163	22.1	1062	13 Q8AXC7	Q8axc7 fugu rubrip
26	1162	22.0	1078	13 Q8AXC8	Q8axc8 fugu rubrip
27	1141.5	21.6	1059	13 Q9DE49	Q9de49 brachydanio
28	1126	21.4	1087	13 Q9PUF6	Q9puf6 gallus gall
29	1111.5	21.1	1097	11 Q8R406	Q8r406 rattus norv
30	1087.5	20.6	1106	4 Q8N5L4	Q8n5l4 homo sapien
31	1080	20.5	1048	13 P79749	P79749 fugu rubrip
32	1079	20.5	1019	13 Q8UVR8	Q8uvr8 fugu rubrip
33	1046	19.8	923	6 Q97745	Q97745 sus scrofa
34	1038	19.7	986	13 Q8UVR9	Q8uvr9 fugu rubrip
35	965.5	18.3	563	11 Q925F7	Q925f7 rattus norv
36	954.5	18.1	1302	13 Q8AXB3	Q8axb3 brachydanio
37	952.5	18.1	1173	13 Q9PTL0	Q9ptl0 brachydanio
38	950.5	18.0	1301	13 Q8UUV9	Q8uu9 brachydanio
39	948.5	18.0	1327	13 Q8QHL3	Q8qhl3 gallus gall
40	934	17.7	1345	11 Q8VCD0	Q8vcd0 mus musculu
41	932	17.7	1379	13 P79701	P79701 coturnix co
42	916.5	17.4	323	11 Q9EQ22	Q9eq22 rattus norv
43	909.5	17.2	1363	11 Q91ZT1	Q91zt1 rattus norv
44	892.5	16.9	323	11 Q9EQ24	Q9eq24 rattus norv
45	841	15.9	824	13 Q90749	Q90749 gallus gall

ALIGNMENTS

RESULT 1					
Q8BR20					
ID	Q8BR20	PRELIMINARY;	PRT;	406 AA.	
AC	Q8BR20;				
DT	01-MAR-2003 (Tremblrel. 23, Created)				
DT	01-MAR-2003 (Tremblrel. 23, Last sequence update)				
DT	01-MAR-2003 (Tremblrel. 23, Last annotation update)				
DE	FMS-like tyrosine kinase 3.				
OS	Mus musculus (Mouse).				
QC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=C57BL/6J; TISSUE=Brain;				
RX	MEDLINE=22354683; PubMed=12466851;				
RA	The FANTOM Consortium,				
RA	The RIKEN Genome Exploration Research Group Phase I & II Team;				
RT	"Analysis of the mouse transcriptome based on functional annotation of				
RT	60,770 full-length cDNAs."				
RL	Nature 420:563-573(2002).				
DR	EMBL; AK045865; BAC32514.1; -				
SQ	SEQUENCE 406 AA; 45834 MW; 155394D1676D7D37 CRC64;				
Query Match	32.4%;	Score	1707.5;	DB 11;	Length 406;
Best Local Similarity	80.0%;	Pred. No.	1.3e-135;		
Matches	324;	Conservative	30;	Mismatches	50;
		Indels	1;	Gaps	1;
Qy	1	MPALA-RDAGTVPLLVVFSAMIFGTTTNDLPVTKVLINHKNDSSVGKSSSYPMVSES	59		
Db	1	MEALAQSRDRELLLVLSVMILETVTNDLPVTKVLISHENNGSSAGKSSYEMVRGS	60		
Qy	60	PEDJGCLARPOSGTVEAAAEVVDVSGASITLQVLVDAPGNISCLWVPKHSNLCQPHFD	119		
Db	61	PEDLQCAPRROSGTVEAAAEVVDVSGASITLQVLVDAPGNISCLWVPKHSNLCQPHFD	120		
Qy	120	LQNRGVSVMLTKTQTCAGYLLFIQSEATNYTILFTVSTNTLLYTLRRPYFRKMQ	179		
Db	121	LQNRGVSVMLTKTQTCAGYLLFIQSEATNYTILFTVSTNTLLYTLRRPYFRKMQ	180		

Qy 180 DALVCISSVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGR 239  
 Db 181 DALLCISGEVPEPTVWVLCSSHRESCKEESPAVVKKEKVLHELFGTDIRCCARNALGR 240  
 Qy 240 ECTRLFTIDNQTPTTLPLQFLKVGELWIRCKAVHVNHGFLTWELNKALEEGNYFE 299  
 Db 241 ECTRLFTIDNQAQOSTLPLQFLKVGELWIRCKAVHVNHGFLTWELNKALEEGNYFE 300  
 Qy 300 MSTYSTNRTMIRILFAFVSSVARNDTGYTSSSKHPSQSALVTIVGKGFINATNSSBDY 359  
 Db 301 MSTYSTNRTMIRILFAFVSSVGRNDTGYTSSSKHPSQSALVTILEKGFINATSSQOEY 360  
 Qy 360 BIDYEEPCFVRKAYPOICTWTTPSRKSPCEQKGLDNGYSIS 404  
 Db 361 BIDPYEKFCSVRKAYPRICRTWTFQSAPFCEQGLEGDGYRLS 405

RESULT 2  
 Q8JFRS PRELIMINARY; PRT; 976 AA.  
 AC Q8JFRS;  
 DT 01-OCT-2002 (TremBLrel. 22, Created)  
 DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)  
 DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)  
 DE SI:d2116L04.1 (kit receptor (Tyrosine kinase)).  
 GN KIT.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OC NCBI\_TaxID=7955;  
 RN 1)  
 RP SEQUENCE FROM N.A.  
 RA Pandian R.;  
 RL Submitted (Aug-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL691516; CAD43458.1; --  
 DR InterPro; IPR003599; Ig.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig c2.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR001824; RtkinaseII.  
 DR InterPro; IPR002290; Ser\_thr\_kinase.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR Pfam; PF00047; Ig\_3.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR SMART; SM00409; IG; 4.  
 DR SMART; SM00408; IGC2; 3.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR SMART; SM00219; TyKc; 1.  
 DR PROSITE; PS00835; IG LIKE; 2.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
 KW ATP-binding; Immunoglobulin domain; Kinase; Receptor; Transferase.  
 SQ SEQUENCE 976 AA; 109293 MW; A56921FA48D58ED CRC64;

Query Match 24.1%; Score 1272.5; DB 13; Length 976;  
 Best Local Similarity 38.5%; Pred. No. 2.8e-98;  
 Matches 294; Conservative 122; Mismatches 262; Indels 85; Gaps 20;

Qy 253 PQTTLPO---LFLKVGELWIRCKAVHVNHGFLTWEL---NKALEEGNYEMETYSYN 306  
 Db 206 PPITLQGPORVLLTQGEKLSLCSSTSNVNSDIKWKAPNGVNPVHQNSHLTLTFPIHV 265  
 Qy 307 RTMIRILFAFVSSVARNDTGYTSSSKHPSOSA---LVTIVGKGFINATN--SSDYEID 362  
 Db 266 RTAILSL-----SSVTQDAGNYSCEAINKEGTAKPWNVIYEKGFINTSDVNSTRRVR 321  
 Qy 363 QYEEFCFSVRKAYPOIRC-TWTFGRKSPFCEQKGLDNGYSISKFCNH----- 409

Db 322 AGESLSRVVMAVYPKPHTFSWSYS-----GVKLTNTDVIHTRTSGNSVT 368  
 Qy 410 -----KHQOGEYIFHAENDDAQFTKMTLIRKRPQVLAESA--SQASCFSDGYP 458  
 Db 369 SELKVLRLKVSSEGIYTFSCILNRDATIRQTEVHVISKPOIVSYEGPDIGOVRCVAGGYP 428  
 Qy 459 LPSWTW-----KKCSDKSPNCTEELTEGVWNRKANRKFVGQWVSSSTLANSEAIKGPL 511  
 Db 429 TPQIKWYICDLPHSRCSNLL-NATQE--EEDVVTITMTNPPFGKGAVERLNIK--NNYA 484  
 Qy 512 VKCAYNSLGTSCETILLNSPGPPF---FIQDNISFVATICGVCLLFIIVLLTILLI CHKYK 568  
 Db 485 TLECVASANGEIVTLFISSENTVPHLFTPLLIGFVAAVILVLLIVLT----YKMQ 540  
 Qy 569 QFRYESQLQWQVQTSSSDNEYFYVDPREYEDLKWEFPRENLEFGKVLGSGAFGKMMNAT 628  
 Db 541 KPKYQIQWKVIE--GIHGNNYVYIDPTQLPYDQHWEPFRDKLRFGLTGLSGAFGKVEAT 598  
 Qy 629 AYGISKTVGSIQVAVKMLKEKADSSEREAALSELKMTQLGSHENIVNLLGACTILSGPIY 688  
 Db 599 AYGMSKADVTMTAVKMLKPSAHATEKALSELKVLKSYLGHNINIVNLLGACTVGGPTL 658  
 Qy 689 LIPEYCCVGDLLNVLRSKREKFRHTWTEIFKEHNFSPYPTFQSHNSMPGSREYQ---- 744  
 Db 659 VITEYCCFGDULNLRRLRRRVFYVT---TLGEDAYYRNVMQSEPNDSRNGYTWTKPSVL 715  
 Qy 745 -IHPDSQISGLHGNSHSEDEIEYENOKRLEEDLNVLTFEDLLCFAYQVAKGMFELE 803  
 Db 716 GILSSENRRLSNGDSYSDSDAV-----SEILOEDGLTLDT-EDLLSFSYQVAKGMDFLA 769  
 Qy 804 PKSCVHRDLAARNVLVTHGKVKVCKDFGLARDIMSDSNVVRGNARLPVKWMAPESLREG 863  
 Db 770 SKNCIHRDLAARNILLTQGRVAKICDFGLARDITTDNSYVVKGNARLPVKWMSPESEIFEC 829  
 Qy 864 IYTIKSDVMSYGILLWEIFSLGVNPGYGPVDPANFYKLIQNGFKMDQDFYATEEYIIMQ 923  
 Db 830 VYTESDVMYSYGILLWEIPSLGSSPYPGMPVDSKFKYMKIEGRVWSEFSPSEYDINH 889  
 Qy 924 SCWAFDSRKRPSFPNLTSLFGLCQLADAEAMYQNVNDRVSECP 966  
 Db 890 SCWDADPVKRPSPFSKIVEKIEQISDSTKHVLYNFSSRLPAAP 932

RESULT 3  
 Q9W755 PRELIMINARY; PRT; 976 AA.  
 AC Q9W755;  
 DT 01-NOV-1999 (TremBLrel. 12, Created)  
 DT 01-NOV-1999 (TremBLrel. 12, Last sequence update)  
 DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)  
 DE Kit receptor tyrosine kinase.  
 GN KIT.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OC NCBI\_TaxID=7955;  
 RN 1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99396707; PubMed=10393121;  
 RA Parichy D.M., Rawls J.F., Pratt S.J., Whitfield T.T., Johnson S.L.;  
 RT "Zebrafish sparse corresponds to an orthologue of c-kit and is  
 RT required for the morphogenesis of a subpopulation of melanocytes, but  
 RT is not essential for hematopoiesis or primordial germ cell  
 RT development.";  
 RL Development 126:3425-3436(1999).  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
 CC PROTEIN KINASES.  
 DR EMBL; AF153446; AAD41890.1; --  
 DR HSSP; P11362; 1FGK.  
 DR ZFIN; ZDB-GENE-980526-464; kit.







```
RESULT 6
Q63116 PRELIMINARY; PRT; 978 AA.
AC DT 01-NOV-1996 (TrEMBLrel. 01, Created)
AC DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE C-kit receptor tyrosine kinase.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=92003944; PubMed=1912577;
RA Tsujimura T., Hirota S., Nomura S., Niwa Y., Yamazaki M., Tono T.,
RA Mori E., Kim H., Kondo K., Nishimune Y., Kitamura Y.;
RT "Characterization of Ws mutant allele of rats: A 12-base deletion in
RT tyrosine kinase domain of c-kit gene.";
RL Blood 78:1942-1946 (1991).
CC -! SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -! SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
DR EMBL; D12524; BAA02094.1; -.
DR HSP; P11362; IFGK.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig c2.
DR InterPro; IPR003006; Ig MHC.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001824; RTKinaseIII.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 2.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00219; tyrc; 1.
DR PROSITE; PS00835; IG LIKE; 3.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW Phosphorylation; Receptor; Transferase; Transmembrane;
KW Tyrosine-protein kinase.
SQ SEQUENCE 978 AA, 109341 MW, 0958C33F19889051 CRC64;

Query Match 24.0%; Score 1263.5; DB 11; Length 978;
Best Local Similarity 33.8%; Pred. No. 1.6e-97;
Matches 334; Conservative 171; Mismatches 361; Indels 121; Gaps 37;

QY 35 CVLLNHKNDSSVGKSSYPWVSESPEGLCALRPOSSGTVVEAAAVEVDVSASITLQVL 94
DB 12 CVLLVLRGQTGSQPSAGPEPSP-----SIQPAQS-----ELIVEAGDIRLT 57
QY 95 VDAPGNISCLWFKHSLNLCQPHFDLQNRGWSMVLKMT-E-TOAGEYLLFIOSEATNYT 153
DB 58 CTDPAFVK--WTFE-----ILDVRIENKQSEWIREKAEATHGKYTCVSGSLRSSI 107
QY 154 ILFTVSIRN-TLLYLRPPYFRKQENODALVCISSEVPPIVWVLCDGSGESCKEESPA 212
DB 108 YVF--VRDPAVLFLVGLPLFGK-EDNDALVRCPLTDPO-VSNYSLETCDGKSLPTDLKF 162
QY 213 V-----VKKEKVLHFGDTRCCARNELGRCETRLFTID----LNQTPQTLTLPOL 260
DB 163 VNPKAGITIKNRYAHLRC-----IRCAAQREGKWRSDKFTLKVRAAIKAIWPVSPET 219
QY 261 --FLKVGEPILWIRKAVHVNHGFLTW-----ELENKALEEGNVFEMS--TYSTNRTMIR 311
DB 220 SHLLKEGDTFTVCTIKDVSSTVDSMWIKLNPOQSKAQVKRNSWHOGDFNYERQETILT- 278
QY 312 ILFAFVSSVARNDTGYTTCSSSKHPQSALVT---IVGKGFN---ATNSSEYVIDQYE 365
```

```
Db 279 -----ISSARVNDSGVFMVCYANNFTGSANVTTLTKVVEKGFNIPVKNVT--VFVTDGE 331
QY 366 EFCFSVRFKAYPO-IRCTWTFSRKSPCEQKGLD-----NGYSISKFCNKH-----QP 413
Db 332 NVDLVVEFEAYPKPHQOQIYNRT--PTNRGDDYVKSNDQSNIRVYNELRLTRKGTG 389
QY 414 GEYIFHAENDDAQFTKMTFLINIRKQVILA--EASASQSCFSDGYPSPSWTKKCSDKS 471
Db 390 GTYTFLVNSDVSASVTFDYYVNTKPEILTYDRLMNGRLQCVAAGPPEPTIDWYFCTGAE 449
QY 472 PNCETBEITE-GVMNRKANRKFQGWSSSTLWNSAIGKFLVKCCAYNSLGTSCETILLN 530
Db 450 QRCTVPVPPVDVQIQNASVSPFKLVQSSIDSSVFRHNGTVECKASNAVGK--SAFEN 507
QY 531 SPGPFPF-----IQDN-----ISFVATIGVCLLFIWLTLLCHYKKGFRYESQL 576
Db 508 -----FAFKNSKEQIQPHLTFLPLIGFVVVTA--LWGIIVMWL--AYKLPQKMTQVQW 559
QY 577 QMV-QVTSSDNEYFVDPREYEDLKWFFRENLEFGKVLGSGAFGKVMNATAYGISKT 635
Db 560 KVEEING---NNYVYIDPTQLPYDHKWEFPNRNLSFGKTLGAGAFGKVVEATAYGLIKS 616
QY 636 GVSIOVAVKMLKEKADSSERREALMSLKMVTLQSHENIVNLLGACTLGSPIYLIFEYCC 695
Db 617 DAAMTVAVKMLKPSAHLTEREALMSLKVLSYLGNNHNVNLLGACTVGGPTLVITEYCC 676
QY 696 YGDLNLVLSKREKFRHTWTEIFKEHNFSPYTFQSHPNSSMPGSR-VQIHDDSDOISG 754
Db 677 YGDLNLFRRKDSF--IFSQEQEADAALYKMLLSKSSCDSSNEYMDMKPGSVYVP 734
QY 755 LHGNSPFS---EDELEYENQKLEEDLNLTFFEDLLCFAYQVAKGMELEFKSCVHRD 811
Db 735 TKTKRRSARIDSYIERDVTVAIMEDDEL-ALDLEDLSFSYQVAKGMAFLASKNCIHRD 793
QY 812 LAARNVLVTHGKVKICDFGLARDIMSDSNVYVVRGNARLPVKWMAPESLFEGYITIKSDV 871
Db 794 LAARNILLTHGRITKICDFGLARDIRNDSYVYVVGKGNARLPVKWMAPESEIENCVTFESDV 853
QY 872 WSYGILLWEIFSLGNVPYGPVDANFYKIQNGFQMDQPFYATEEIIYIMQSCWAFDSR 931
Db 854 WSYGIFLWELFSLGSPYGPMPVDKFKYKMKKEGFRMLSPHAPAAAMYEMTKCWDADPL 913
QY 932 KPSPFNLTSLFGCQLADAEAEAMQNV 958
Db 914 KRPTFKQVVQLIEKQISDSSKHIYSNL 940

RESULT 7
O97744 PRELIMINARY; PRT; 964 AA.
AC O97744;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Mast/stem cell growth factor receptor (Fragment).
GN KITI*0101.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Hampshire breed; TISSUE=Leukocyte;
RX MEDLINE=98391767; PubMed=9724328;
RA Marklund S., Kijas J., Rodriguez-Martinez H., Roennstrand L., Funa K.,
RA Moller M., Edfors-Lilja I., Andersson L.;
RT "Molecular evolution of the dominant white phenotype in the domestic
RT pig.";
RL Genome Res. 8:826-833 (1998).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=Hampshire breed; TISSUE=Leukocyte;
RA Marklund S.;
```

RL Thesis (1997), Department of Animal Breeding and Genetics.  
 RL Swedish University of Agricultural Sciences, Uppsala, SWEDEN.  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
 CC PROTEIN KINASES.  
 DR EMBL: AJ223228; CAA1196.1; -.  
 DR HSSP: P11362; IFGK.  
 DR InterPro: IPR003599; Ig.  
 DR InterPro: IPR007110; Ig-like.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR InterPro: IPR001824; RTKinaseIII.  
 DR InterPro: IPR001245; Tyr\_kinase.  
 DR Pfam: PF00047; Ig; 1.  
 DR Pfam: PF00069; pkinase; 1.  
 DR ProDom: PD000001; Prot\_kinase; 2.  
 DR SMART: SM00409; IG; 3.  
 DR SMART: SM00219; TyrcK; 1.  
 DR PROSITE: PS00835; IG LIKE; 3.  
 DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE: PS00011; PROTEIN KINASE DOM; 1.  
 DR PROSITE: PS00109; PROTEIN KINASE TYR; 1.  
 DR PROSITE: PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
 DR ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;  
 KW Phosphorylation; Receptor; Transferase; Transmembrane;  
 KW Tyrosine-protein kinase.  
 FT NON TER 964 964  
 SQ SEQUENCE 964 AA; 108287 MW; 90E72EDFAB1358B7 CRC64;

Query Match 23.9%; Score 1262.5; DB 6; Length 964;  
 Best Local Similarity 33.4%; Pred. No. 1.9e-97;  
 Matches 332; Conservative 169; Mismatches 349; Indels 143; Gaps 38;

QY 47 VGSKSSVPMVSESPEDLGCAIRPQSSGTVEAAAVEVDVSATITLOLVDPAGNISCLWV 106  
 DB 20 VQTGSSQPSV--SPEEL-----SPSHPAKS-ELIVSAGDEIRLCTDPDSVK--WT 67

QY 107 FKHSLNCOFHPDLQNRGVVSMVLKMTETQAGEYLLFIOSEATNYTLFTVSRN-TLL 165  
 DB 68 FETLG-----QLSENTHAEWIVEKAEAMNTGYTCTNEGGLSSIIYVE---VRDPEKL 117

QY 166 YTLRRYFRKMNQDALVCISEVSPPIVEWVLCDSQGESCKEESPAV-----VKKE 217  
 DB 118 FLVDPPLYGK-EDNALVRCLPTDPE-VTNYSLTGCCEKPLKDLTFVADPKAGITIRNV 175

QY 218 EKVHLELFGTDIRC--CARNELGRE-CTRLFTID---LNQTFQTLTPQ--LFLKVGEP 268  
 DB 176 KREYHRL-----CLHCSANQGGKSVLSKFTLKRAAIRAVPVAVSKASYLLREGEEF 229

QY 269 WIRCKAVHVNHGFLTWELN---KALEEGNYFEMSTYSTNRTMIRILFAFVSSVARNDT 325  
 DB 230 AVNCLIKDVSSSDSMWIRENSQTKAQVKNRSHQGDEN---FLRQERLTISSARVND 285

QY 326 GYTCSSSKHPSQSALVT-----IVKGFINA-----TNSSEDEVIDQVEEFCFSVRP 373  
 DB 286 GVPWCYANNITFGSANVTTLLEVDKGFINIFPMNNTTVFVNDGEDVDL-----IVEY 337

QY 374 KAYPO-----IRCTWTSRKSFPCEQKGLDNGYSISKFCNKH-----OPGEYIF 418  
 DB 338 EAYPKPEHRQWIIYNNRTATDKWEDYPKSE-----NESNIRYSELHLTRLKGTGGTYTF 392

QY 419 HAENDDAQFTKMTLNRIRKPVQLA--EASASQASCFSDGYPPLPSMTWKKCSKSPNCTE 476  
 DB 393 LYSNADVNSVTENVYNTKPEILTHDRLNMGMLQCVAAAGFPPTIDWYFCPTEQRCSV 452

QY 477 EI-TEGVNMRKANRKYFGQWSSSTLMSBAIKFLVKCCAYNSLGTSCETILLNSPGPF 535  
 DB 453 PVGPDVQIQNSVSPGKLVIHSSIDYSFAFKINGTVECRAYNDVGKS--SAPPN-----F 506

QY 536 PFTQD-----NISFATTGVCLLFFIVLTLLICHKKYKKQFRYESQLQWV-QVTGS 584  
 DB 507 AFKEQIHAHTLFTLLIGFVIAAGMCMCIIVMLT---YKLYQPMYEVQWVKEEING- 561

QY 585 SDNEYFYVDPREVEYDLKWEFFRENLEFGKVLGSGAGFKVMNATAYGISKTGYIOVAVK 644  
 DB 562 --NNVYIIDPTQLPYDHKWEFFRNRLSPFKTLGAGAFGKVEATATGLIKSDRAMTVAVK 619

QY 645 MLKEKADSSERELMSSELKQMTQLGSHENIVNLLGACTLSGPIYLIPIFYCCYGDLLNLYR 704  
 DB 620 MLKPSAHLTEREALMSSELKLVSYLGNHMINVLLGACTTGGPTLVITEYCCYGDLLNLYR 679

QY 705 SKREKPHRTWTE-----IPKEHNFSTFYTFQSHPNSSM---PGSREVOIHPSDDQISGL 755  
 DB 680 RKRDSPICQKQEDHAEALYKNLLHSKSSCSDSTNEYMDKPGSVYV-VPTKADKERSA 738

QY 756 HGNSFHSEDEIEYENQKRLLEEEDNLVLTFCFAYQVAKGMEFLFKSCVHRDLAAR 815  
 DB 739 RIGSY-----IERDVTPTAMEDEL-ALDLELLSFSYQVAKGMFLASNCIHRDLAAR 792

QY 816 NVLVTGKVVKICDFGLARDIMSDSNVYVRGNARLPVKWMAPESLFEGIYTIKSDVMSYG 875  
 DB 793 NILLTHGRITKICDFGLARDIKNDSNVYVKGARLPVKWMAPESFVNCVYTFESDWSYG 852

QY 876 ILLWEIFSLGVNYPGIPVDANFYKLIQNGKMDQPFYATEEYIIMQSCWAFDSRKRPS 935  
 DB 853 IFLWELFSLGSPYPCMPVDSPFYKMKEGFRMLSPHAPAEYMDIMKTCWDADPLKRPT 912

QY 936 FPNLTISFLGCOLADAEAMYNQVDRVSEC-PH 967  
 DB 913 FKQIVQLIEKQISESTNHIYSN---LANCSPH 941

RESULT 8  
 Q8AXC6 PRELIMINARY; PRT; 984 AA.  
 AC Q8AXC6;  
 DT 01-MAR-2003 (T-EMBLrel. 23, Created)  
 DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)  
 DE Kinase receptor c-kit.  
 GN C-KIT.  
 OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Naeleleostei;  
 OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;  
 OC Tetraodontidae; Tetraodontidae; Takifugu.  
 OX NCBI\_TaxID=31033;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Williams H., Brenner S., Venkatesh B.;  
 RT "Characterization of the Platelet-derived Growth Factor Receptor Alpha  
 and c-kit Genes in the Pufferfish Fugu rubripes.";  
 RL DNA Seq. 13:263-270(2002).  
 DR EMBL; AF456419; AAN87555.1; -;  
 KW Receptor.  
 SQ SEQUENCE 984 AA; 110732 MW; 27D1EDC837CE8295 CRC64;

Query Match 23.9%; Score 1262; DB 13; Length 984;  
 Best Local Similarity 38.2%; Pred. No. 2.2e-97;  
 Matches 290; Conservative 114; Mismatches 255; Indels 100; Gaps 21;

QY 266 EPLWIRCKAVHVNHGFLTWEL---ENKALEEGNYFEMSTYSTNRTMIRILFAFVSSVAR 322  
 DB 220 ESYLTLCTNTVNGNTIKLKWAPLGSQPAKVDGS---SRILTFTQARSATLHIAAVRI 276

QY 323 NDTGYTCSSSKH---PSQSALVTIVGKGF1--NATNSSDEYIDQVEEFCFSVRKAYP 377  
 DB 277 QDTGRVQCEAENEKGVSTQSVLDVFEKGFWSNPVNG-TIQVRAGESILLVSIEAYP 335

QY 378 QIR-CTWTSRKSFPCEQKGLDN-----GYSIS---KFCNHK-HQGEYI PHAE 421  
 DB 336 MPASASWSFM-----GRGLHNTSDHVTTRSHVTSSELKLVRLKMEGGVYTFQAS 388

QY 422 NDDAQFTKMTLNRIRKPVQLAEASA--SQASCFSDGYPPLPSMTWKKCSKSPNCTEIT 479  
 DB 389 NGDASVNHHTTIFVISKPEIVSHEGPDVGQVRCVAEGFPAPQITWYCYEQFARCSQVN 448



```

Db 680 RKDSFSCQEDHAEALYKNLLHKSCESSDSTNEMDMKPGSVY-VPTKADKERSA 738
QY 756 HGNSEHSEDEIEYENQKRLSEEDLNVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAAR 815
Db 739 RIGSY-----IERDVTAPINMEDDEL-ALDLEDLLSFYSYQVAKGNAFKSNCHIRDLAAR 792
QY 816 NVLVTGKVKIKIDFGIARDIMSDSNVYVGNARLPVKWMAPELSFEGYIYIKSDVMSYG 875
Db 793 NILTHGRITKICDFGLARDIKNDSNVVVGKGNARLPVKWMAPELSFNCVYTFESDVMSYG 852
QY 876 ILLWEIFSLGVNYPGIPVDANFYKLIQNGFKMDQPFYATEEIIYIMQSCWAFDSRKPS 935
Db 853 IFLWELFSLGSSYPGMPVDSKFYKMKKEGRMLSPHAPVEMYDINKTCDWADPLKRP 912
QY 936 FPNLTSLGCOLADAEABEAMQNDVRVSEC-PH 967
Db 913 FKQIVOLIEKQISESTNHIYSN-----LANCSPH 941

RESULT 10
Q9TQ00
ID Q9TQ00 PRELIMINARY; PRT; 964 AA.
AC Q9TQ00;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Mast/stem cell growth factor receptor (Fragment).
GN KIT1-0202.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Yorkshire / Landrace cross; TISSUE=Leukocyte;
RA Marklund S.;
RL Thesis (1997), Department of Animal Breeding and Genetics,
RL Swedish University of Agricultural Sciences, Uppsala, SWEDEN.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Yorkshire / Landrace cross; TISSUE=Leukocyte;
RX MEDLINE=98391767; PubMed=9724328;
RA Marklund S., Kijas J., Rodriguez-Martinez H., Roennstrand L., Funa K.,
RA Moller M., Edfors-Lilja I., Andersson L.;
RT "Molecular evolution of the dominant white phenotype in the domestic
RT pig.";
RL Genome Res. 8:826-833(1998).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
DR EMBL; AJ223230; CAA11198.1; -.
DR HSP; P11362; IFGK.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001824; RTKinaseII.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 2.
DR SMART; SM00409; IG; 3.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00835; IG_LIKE; 3.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW Phosphorylation; Receptor;
KW Tyrosine-protein kinase.
FT NON_TER . 964

```

```

SQ SEQUENCE 964 AA; 108287 MW; 1B21292A962E9191 CRC64;
Query Match 23.9%; Score 1261.5; DB 6; Length 964;
Best Local Similarity 33.4%; Pred. No. 2.3e-97;
Matches 332; Conservative 169; Mismatches 349; Indels 143; Gaps 38;
QY 47 VGSSSYPWVSSPESPDLCALRPQSSGTVYEAARAVEVDVSIATLQVLVDAPGNISCLWV 106
Db 20 VQTGSSQPSV--SPEEL-----SPPSIHPAKS-ELIVSAGDEIRLFTDPSGVK--WT 67
QY 107 FKHSLSNCPHFDLQNRGVVSVILKMTETQACEYLLFIQSEATNVTILFTVSRN--TLL 165
Db 68 FETLG-----QLSENTHAEWIVEKAEAMVNTGYCTNEGGSSSIYVF--VRDPEKL 117
QY 166 YTLRRYPFRKMENQDALVCISESPBPPIVEMVLCDSQGESCKEESPAV-----VKKE 217
Db 118 FLVDPLYGK-EDNDALVRCPLTDPE-VTNYSLTGCEGKPLPKDLTFVADPKAGITIRNV 175
QY 218 EKVLHELFGTDIRC--CARNELORE-CTRLFTTD----LNQTFQTLTLPQ--LFLKVGCEPL 268
Db 176 KREYHRL-----CLHCSANOGGKSVLSKFTLLKVRAAIRAVPVAVSKASYLLREGEF 229
QY 269 WIRCKAVHNHGFGLTWELN---KALEEGNYFEMSTYSTNRTMIRILFAFVSSVARNDT 325
Db 230 AVNCLIKDVSSVDSMWIRENSQTKAQVRKNSWHQGDNF----FLROEKLITISARVND 285
QY 326 GYITCSSSSKHPQSOSALVT---IVGKGFINA-----TNSSEYEDIDQYEEFCFSVR 373
Db 286 GVPMCYANNTFGSANVTITLLEVVDKGFINIFPMWNTTVFVNDGEDVDL-----IVEY 337
QY 374 KAYPQ-----IRCTWTSRKSFPCEQKGLONGYSISKCNHKKH-----QPGYIF 418
Db 338 EAYPKPEHRQWIYMNRTATDKWEDYPKE-----NESNIRYVSELHLTLKGTGEGTYTF 392
QY 419 HAENDDAQFTKMETLNIIRKPOVLA--EASASQASCFSDGYPPLSPSTWKKSDKSPNCTE 476
Db 393 LVSNDVNSSVTFNVYNTKPEILTHRLMNGMLQCVAAGFPEPTIDWYFCPTGEQRCSV 452
QY 477 EI-TEGVNRKANRKFVGOWVSSSTLNMSBAIKGFLVKCCAYNSLGTSCETILNPSGPF 535
Db 453 PUGPVDVQIONSVSPFGKLVHSSIDYSAFKNGTVECRAYNDVGKS--SAFEN----F 506
QY 536 PFTQD-----NISFYATTGVCLLFIVLITLICHYKKQFQRYESQLQWV-QVTGS 584
Db 507 AFKEQIHAHTLFTPLLLIGFVIAAGMCIIVMLT---YKYLQKPMYEVQWKVVEEING- 561
QY 585 SDNEYFVDPREYEDLKWEFPRENLEFGVLGSGAGFKYMNATAYGISKTGVSIQVAVK 644
Db 562 --NNYVIIDPTQLPYDHKWEFPNRNLSFGKTLAGAGFKGVVEATAYGLIKSDAAMTAVK 619
QY 645 MLKEKADSSEREALMSELKMTQLGSHENIVNLLGACTLSGPIYLIPEYCCYGLLNYLR 704
Db 620 MLKPSAHLTEREALMSELKVLVSLGNHNVNLLGACTIGGPTLVITVEYCCYGLDLNFLR 679
QY 705 SKREKTHRTWTE-----IPKEHNFSPYPTFQSHPNSSM---PGSRREVQIHPDSQISGL 755
Db 680 RKDSFSCQEDHAEALYKNLLHKSCESSDSTNEMDMKPGSVY-VPTKADKERSA 738
QY 756 HGNSEHSEDEIEYENQKRLSEEDLNVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAAR 815
Db 739 RIGSY-----IERDVTAPINMEDDEL-ALDLEDLLSFYSYQVAKGNAFKSNCHIRDLAAR 792
QY 816 NVLVTGKVKIKIDFGIARDIMSDSNVYVGNARLPVKWMAPELSFEGYIYIKSDVMSYG 875
Db 793 NILTHGRITKICDFGLARDIKNDSNVVVGKGNARLPVKWMAPELSFNCVYTFESDVMSYG 852
QY 876 ILLWEIFSLGVNYPGIPVDANFYKLIQNGFKMDQPFYATEEIIYIMQSCWAFDSRKPS 935
Db 853 IFLWELFSLGSSYPGMPVDSKFYKMKKEGRMLSPHAPVEMYDINKTCDWADPLKRP 912
QY 936 FPNLTSLGCOLADAEABEAMQNDVRVSEC-PH 967
Db 913 FKQIVOLIEKQISESTNHIYSN-----LANCSPH 941

```

RESULT 11  
Q8WN23 ID Q8WN23 PRELIMINARY; PRT; 979 AA.  
AC Q8WN23, 20, Created)  
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)  
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)  
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
DE C-KIT.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Zemde D., Yuzbasiyan-Gurkan V.,  
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
CC -|- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).  
CC -|- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
CC PROTEIN KINASES.  
DR EMBL: AF448148; AAL40833.1; --  
DR InterPro; IPR003599; Ig-like.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003006; Ig MHC.  
DR InterPro; IPR000719; Prot Kinase.  
DR InterPro; IPR001824; RTKinaseIII.  
DR InterPro; IPR001245; RTKinase.  
DR Pfam; PF00047; Ig; 2.  
DR Pfam; PF00069; pkinase; 1.  
DR ProDom; PD000001; Prot\_kinase; 2.  
DR SMART; SM00409; IG\_3.  
DR SMART; SM00219; TyKc; 1.  
DR PROSITE; PS00835; IG-LIKE; 2.  
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.  
DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.  
DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;  
KW Phosphorylation; Receptor; Transmembrane;  
KW Tyrosine-protein kinase.  
SQ SEQUENCE 979 AA; 109753 MW; 46C30D5DB8E33D3 CRC64;

Query Match 23.78; Score 1250.5; DB 6; Length 979;  
Best Local Similarity 33.68; Pred. No. 2e-96;  
Matches 334; Conservative 161; Mismatches 345; Indels 153; Gaps 39;  
QY 47 VGGSSYPVMSPEPDLGALRPOSSGTVYEAFAVEVDVSASITLQVLVDAPGNISCLWV 106  
DB 22 VQTGSSQPSVSGEPLSP-SIHPAKS-----ELIVSGDELRLSCTDPGFVK--WT 69  
QY 107 FKH-SLNCOPHDLQNRGVVSMVLKMTQTQAGEYLLFIQSEATNYTIL-----FT 157  
DB 70 FETLGQLNENTH---NEWITE-----KAEG-----HTGNYTCTNRDGLSRISY 110  
QY 158 VSRN-TLLYTLRRPRKMNQDALVCISGPEPIVWVLCDSQGESKESPAV---213  
DB 111 VFVRDPAKFLVDLPYLGK-EGNDTLVRCPLTDP-E-VTNSLRGCEGKPLKDLTFVADP 168  
QY 214 ----VKCEKVLHFGTDIRC--CARNELGRE-CTRLFTID-----LNQTPQTTLPLQ- 260  
DB 169 KAGITRNKVEYHRL-----CLHCSADQGRVLSKFTLKVRAAIRAVPVVSVSKTS 222  
QY 261 -FLKVGELPWRKCAVHNVHFGTLTWELEN-----KALEEGNYFEMSTYNTNRWIRLFA 315  
DB 223 SLLKEGEAFVSMCFIKDVSSFVDSMIKENSQQTNAQTQSNWHHGDFNFERQEKLI--- 279  
QY 316 FVSSVARNDTGYTCSSKHPSQALVT-----IVGKGFINA-----TNSSEDEIDQ 363  
DB 280 -ISSARVNDSGVPMCVANNITFGSANVTTLLEVVVDKGFINIFPMWSTITFVNDGENDVL-- 336  
QY 364 YBEFCFSVFKAQYQ-IRCTWTSTRKSPFCEQKGL---DNGYISIKFCNHHK-----Q 412

337 -----IVEYEAYPKPEHQWIMYMRFTFTDKWEDYPKSDNESNI-RYVSELHLTRLKGN 369  
QY 413 PGEYIFHAENDDAQFTKMTNLIRKPKQVLAESASQA--SCFSDGYPLPSTWTKKCSK 470  
DB 390 GGTFTFQVNSDVSSTVYVNTKPEILTHESLTNGMLQCWAGFPPEPVDWYFCPGA 449  
QY 471 SPNCTEEI-TEGVNKRKANRVFGQWSSSTLNMSEAIGPLVKCAVNSLGTSCETILL 529  
DB 450 EQRCSVPFGPMDVQMONSSLSFGKLVVQSSIDYSAFKHNGIVECRAYNNVGRS--SAFF 507  
QY 530 NSPGPPFPFIQDN-----ISFYATIGVCLLFIIVVLTLLIICHKKYKQFYESQ 575  
DB 508 N----PAFKGNSKEQIHPHTLFTPLLIGFVIAAGMCIIVMLT----YKYLQKPMYEVQ 559  
QY 576 LQMV-QVTGSSDNEYFYVDREYEDLKWEPRENLEFGKVLGSGAFGKVMNATAYGISK 634  
DB 560 WKVVEEING---NNYVYIDPTQLPYDHKEWEPFRNRLSPGKTLGAGAFGKVVETAYGLIK 616  
QY 635 TGVSIQVAVKMLKEKADSSEREAELMSELKMTOLGSHENIVNLLGACTLSGGPIYLIFEYC 694  
DB 617 SDAANTVAVKMLKPSAHLTEREALMSELKVSILYGNHWNIVNLLGACTVGGPTLVITEYC 676  
QY 695 CYGDLNLYRSKREKFRHTWTE-----IFKEHNFSFYPTFQSHPNSSM---PGSREV01 745  
DB 677 CYGDLNLFRLKRDSDFCCKQEDHGEVALYKNLLHKSSESCSDSTNEYMDKPGVSYV-V 735  
QY 746 HPDSQISGLHNSFHSEDEIEYENQKRLBEEEDLNVLTTFEDLLCFAYQVAKGMFLPEK 805  
DB 736 PTKADKRSARIGSY-----IERDVTPTAIMEDDEL-ALDLEDLLSFSYQVAKGMFLASK 789  
QY 806 SCVHRDLAARNVLVTHGKVKVICKDPLGLARDTMSDSNYVVRGNARLPVKWMAPESEFEGY 865  
DB 790 NCHRDLAARNILLTHGRITKICDPLGLARDIKNSNYVVRGNARLPVKWMAPESEFENCVY 849  
QY 866 TIKSDVMSYIGILLWEIFSLGVPYGPVDANFYKLIQNGFKMQDPFVATEEIIYIMQSC 925  
DB 850 TPESDWSYIGIFLWELSLGSSPYGMPVDSKFTKMKKEGFRMLSPHAPAEYDIMKTC 909  
QY 926 WAFDSRKRPSFPNLTSLFGCOLADAEAMQNV 958  
DB 910 WDADPLKRPFTKQIVQLIEKQISDSTNHIYSL 942  
RESULT 12  
Q99662 ID Q99662 PRELIMINARY; PRT; 972 AA.  
AC Q99662; 03, Created)  
DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)  
DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)  
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
DE KIT protein.  
GN KIT.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92228497; PubMed=1373482;  
RA Andre C., Martin E., Cornu F., Hu W.X., Wang X.P., Galibert F.;  
RT "Genomic organization of the human c-kit gene: evolution of the  
RT receptor tyrosine kinase subclass III.",  
RL Oncogene 7:685-691(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97179223; PubMed=9027509;  
RA Andre C., Hampe A., Lachaume P., Martin E., Wang X.P., Manus V.,  
RA Hu W.X., Galibert F.;  
RT "Sequence analysis of two genomic regions containing the KIT and the  
RT FMS receptor tyrosine kinase genes.",  
RL Genomics 39:216-226(1997).  
CC -|- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).  
CC -|- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-

CC PROTEIN KINASES.  
 DR EMBL; U63834; AAC50969.1; --  
 DR HSSP; P11362; 1FGK.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig\_c2.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR001824; RTKinaseIII.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00047; Ig; 1.  
 DR Pfam; PF00069; pkinase; 1.  
 DR ProDom; PD000001; Prot\_kinase; 2.  
 DR SMART; SM00408; IGC2; 1.  
 DR SMART; SM00219; TyrKc; 1.  
 DR PROSITE; PS08335; IG LIKE; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00109; PROTEIN KINASE\_TYR; 1.  
 DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
 KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;  
 KW Phosphorylation; Receptor; Transmembrane;  
 KW Tyrosine-protein kinase.  
 SQ SEQUENCE 972 AA; 109450 MW; D59DEP9AF761FDA CRC64;

Query Match 23.6%; Score 1247; DB 4; Length 972;  
 Best Local Similarity 32.8%; Pred. No. 4e-96;  
 Matches 323; Conservative 166; Mismatches 353; Indels 144; Gaps 34;

QY 47 VGKSSYPVMSSEPDGALRQSSQGVYEAADVDSASITLQVLVDAPGNISLWV 106  
 DB 20 VQTGSSQPSVSPG-EPSPSIHFGKSDLI-----VRGDEIRLLCTDTPGVK--WT 67

QY 107 FHSSSLNCPHFLDQNRGVVSMILKMTQAGEYLLFIQSEATN---TYLFTVTSIRNT 163  
 DB 68 FE-----ILDETENKQEWIT-EKAEATNTGKYCTCNKHGLSNS 106

QY 164 L-----LYTLRRPYFRKMNQDALVCISSEVPPIVEWVLCDSQGSKKEE----- 209  
 DB 107 IYVFVRDPAPKFLVDRSLYK-EDNDTLVRCPLTDPE-VTNYSLKGGKRLPKDLRFIP 164

QY 210 ---SPAVVKEEVKLHFLGTDIRCCARNELGRECTRLFTIDL---NQTPQTLTPQ--L 260  
 DB 165 DPKAGIMIKSVKRAYHRLC---LHCSVDQBGKSVLSEKFLKVPAPKAVPVVSVSKASY 221

QY 261 FLVKVGLPLWIRKAVHVNHGFLGTWELN---KALEEGNYFEMSTYSTNRTMIRLPFV 317  
 DB 222 LAREGEFTVCTIKOVSSSVYSTWKEHSQTKLQEKYNSWHGDFNYERQAT-----LTI 277

QY 318 SSVARNDTGYTSCSSKHPQSALVT---IVGKGFNA-----TNSSEYDEIDQYE 365  
 DB 278 SSARVNDSGVFMVCYANNTFGSANVTTLLEVVDKGFNIFFPMINTVFVNDGENVDL---- 333

QY 366 EFCFSVRFKAYPQ-IRCTWTFSRKSF--PCEQKLDNGYSISKPCNHK-----QPGE 415  
 DB 334 ---IVEYEAPPKEHQOYIMNRTFTDKWEDYPKSENEIRVSELHLTRLKGTGGT 389

QY 416 YIFHAENDDAQTKMFTLNIRRKQVLA--EASASQASCFSDGYPLPSWTWKCKSDKSPN 473  
 DB 390 YTLVNSVDVNAALAFVNVYNTKPEILTYDLVNGMLQCVAAGPEPTIDWYFCPTGEQR 449

QY 474 CTEITE-GVNNRKANKRVQGVSSSTLNNSEAIKGLVKCCAYNSLGTSCETILNLSNP 532  
 DB 450 CSASVLVDVQTLNSSGPPFGKLVQSSIDSSAFKHGTVCKEAYNDVGKT--SAYFN-- 505

QY 533 GPFPFIQD-----NTSFYATIGVCLLFIVVLTLLCHKYKQFRYESQLQMW-QV 581  
 DB 506 --FAFKEQIHPHTLFTPLLLGFVIVAGMCIIVMLT-----YKYLQPMYEVQWKVVEI 559

QY 582 TGSDDNEFYVDREYEDLKWEPPRENLEFGKVLGSGAFGKVNNAATAYGISTKGVSIQV 641  
 DB 560 NG---NNYVIDPTQLYDHNKEFPRLNLSFGKTLGAGAFGKVNNAATAYGLIKSDAMTV 616

QY 642 AVKMLKEKADSSERREALMSELKMMTQLGSHENIVNLLGACTLSGFIYLIFEYCCYGDLLN 701

DB 617 AVKMLKPSAHLTEREALMSELKVLVSGNHNIVNLLGACTIGTFLVITEYCCYGDLLN 676  
 QY 702 YLRSKEKEFHRTWTE-----IPKEHNFQFYPTFQSHNSSM---PGSRVQIHPDSQDI 752  
 DB 677 FLRRKEDSFICSKQEDHAEALYKNLLHKSBCSDSTNEYMDMKPGSVYV-VPTKADKR 735  
 QY 753 SGLHGNFSHSEDEIEYENQKRLEEBEDLVNLTPEDLICFAYQVAKGMEFLFKSCVHRDL 812  
 DB 736 RSVRIGSY-----IERDVTVPAINMEDDEL-ALDLEDLLSPSYQVAKGMAFLASKNCIHRDL 789

QY 813 AARNVLVTHGKVKICDFGLARDIMSDSNVVRGNARLPVKWMAPESLFPGIYTIKSDVW 872  
 DB 790 AARNVLVTHGRITKICDFGLARDIKDNSYVVKGNARLPVKWMAPESLFNCVYTFESDVW 849

QY 873 SYGILLWEIFSLGWNYPGIPVDANFYKLIQNGFKMDQPFYATEEIIYIMQSCWAFDSRK 932  
 DB 850 SYGIFLWELFSLGSSPYGMPVDSKFKYKMEKGFRLMSPEHAPAEYDYMKTCTWDADPLK 909

QY 933 RPSFPLNLTSLFGQLADAEEMYNQV 958  
 DB 910 RPTFKQIVQLIEKQISESTNHIYSNL 935

## RESULT 13

Q9XS93 PRELIMINARY; PRT; 978 AA.  
 AC Q9XS93;  
 DT 01-NOV-1999 (T-EMBLrel. 12, Created)  
 DT 01-NOV-1999 (T-EMBLrel. 12, Last sequence update)  
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)  
 DE KIT.  
 GN C-KIT.  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 CX NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Cerebellum;  
 RA London C.A., Galli S.J., Yuuki T., Hu Z.-Q., Helfand S.C.,  
 RA Geisler E.N.;  
 RT "Spontaneous canine mast cell tumors express tandem duplications in  
 RT the proto-oncogene c-kit.";  
 RL Exp. Hematol. 0:0-0(1999).  
 CC -I- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).  
 CC -I- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
 CC PROTEIN KINASES.  
 DR EMBL; AF099030; AAD28369.1; --  
 DR HSSP; P11362; 1FGK.  
 DR InterPro; IPR003599; Ig.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR001824; RTKinaseIII.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00047; Ig; 2.  
 DR Pfam; PF00069; pkinase; 1.  
 DR ProDom; PD000001; Prot\_kinase; 2.  
 DR SMART; SM00409; IG; 3.  
 DR SMART; SM00219; TyrKc; 1.  
 DR PROSITE; PS08335; IG LIKE; 2.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00011; PROTEIN KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN KINASE\_TYR; 1.  
 DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
 KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;  
 KW Phosphorylation; Receptor; Transmembrane;  
 KW Tyrosine-protein kinase.  
 SQ SEQUENCE 978 AA; 109651 MW; 5AC31ESAC4E9910F CRC64;

Query Match 23.6%; Score 1244; DB 6; Length 978;  
 Best Local Similarity 33.2%; Pred. No. 7.2e-96;



Matches 330; Conservative 161; Mismatches 346; Indels 158; Gaps 38;	
QY	47 VKSSSYPMVSSPEDLGCALRQSSGTYVEAAAEVDVSASITLQVLVDAPGNISCLWV 106
Db	22 VQTGSQPSVSGEPLP-SIIPAKS-----ELIVSGDELRLSTCTDGFVK--WT 69
QY	107 FKX-SSLNCQPHFDLQNRGVSMVLKMTTETAGAYLLFIQSEATNYTIL-----FT 157
Db	70 FETGLQNLNTH-----NEWITE-----KAEAG-----HTGNYTCTNRDGLRSY 110
QY	158 VSIRN-TLTYLRRPYFRKQENODALVCISESVPEPIVEWVLCDQSGESCKEESPAV--- 213
Db	111 VEVROPALFLVDLPLYGR-EGNDTLVRCLTDP-E-VTNYSLRGCEGKPLPKDLTVPADP 168
QY	214 -----VKKEEVLHFGTDIRC--CARNELGRE-CTRLFTID-----LQNTPTTLQPL- 260
Db	169 KAGIIRNVKRYHRL-----CLHCSADQKRTVLSKRFTLKVRAAIRAVPVVSVKTS 222
QY	261 -FLKVGEPILWIRCKAVHNVHGFGLTWELN-----KALEEGNYFEMSTYSTNRTMIRILPAP 316
Db	223 SLLKEGAFSVMCFIKVSSFFVDSMWIKENSQNAQTQSNWHHGDFNFERQEKLI---- 278
QY	317 VSSVARNTGYTCSKSHPSQSALVT-----IVKGFINA-----TNSSEDEYIDQY 364
Db	279 ISSARVNDSGVPMCVANNTFGSANVTTLLEVVDKGFINIPFMWSTTIPVNDGENVDL--- 335
QY	365 BEFCFSVRFKAYPO-IRCTWTFSRSPFCEQKL-----DNGYSISKPCNHK-----QP 413
Db	336 -----IVEYEAYPEHQOQWIMYNNRTFTDKWDYPKSDNESNI-RYVSELHLTRLKNGEG 389
QY	414 GEYIHAENDDAQTKMFTLNRRRPOVLAEASQA---SCFSDGYPLPSWTWKCSDKS 471
Db	390 GTYTFQVNSDVNSVTNVYNTPEILTHESLNGMLQCVAGPEPAVDWYFCPGA 449
QY	472 PNCTEEI-TEGVNVRKANRKFQGVSSSTLNMSEAIGFLVKCCAYNSLGTSCETILN 530
Db	450 QRCSPVIGMDVMQNSLSLSPGKLVQSSIDYSAPKNGTVECRAYNVGSR--SAPN 507
QY	531 SPGPFPFTQDN-----ISFVATTGVCLLFLVILVTLTLLCHYKQFYRESOL 576
Db	508 -----FAFKGSEQIHPHTLFTPLLIGFVIAAGMMCIIVMLT-----KYLOKPMYEVQW 559
QY	577 QMV-OVTGSSDNEYFYVDFREYDLEKWFRENLEFGKVLGSGAGPKVMNATYGISKT 635
Db	560 KVEEING-----NNYVIIDTQDPYDHKEFPFRNLSPFGKTLGAGAPGVVEATAYGLIKS 616
QY	636 GVSIOVAVKMLKEKADSEREALMSBELKMTQLGSHENIVNLLGACTLSGGTYLIFYECC 695
Db	617 DAAMTVAVKMLKPSAHLTEREALMSBELKVLVSLGNHMINVLLGACTVGGPTLVITEYCC 676
QY	696 YGDLNLNLRKREKPHRTWTETPFKEH-NPSFYPTFQSHNPSMPSGSRVQIHPDSQISG 754
Db	677 YGDLNLFRLRRKRDSPICSKQE---DHGEVALYKNLLHSEKSCSDS-----TNEYMDM 726
QY	755 LHGSHFSEDELEYENQKLEB-----EEDLANVLTPEDLILCFAYQVAKGMEFL 803
Db	727 KPGSYVYPTKADRRSRIGSYIERDVTPAIMEDELALDELLELSPSYQVAKGMFLA 786
QY	804 FKSCVHRDLAARNLVTHGKVVKI CDFGGLARDIMSDSNVVRGNARLPVKWMAPESLPEG 863
Db	787 SKNCIHRDLAARNLITLHTRITKICDFGLARDIKDSDNVVKGARLPVKWMAPESEFNC 846
QY	864 IYTKSDVMSYGILLWEIFSLGVNYPGIPVDANFYKLQNGFKMDQDPYATEEYIIMQ 923
Db	847 VYTPESDVMYSYGIFLWELFSLGSSPYPGMPVDSFKYKMIKEGFRMLSPEHAPEMYDIMK 906
QY	924 SCWAFDSKRSFPNLTSLFGLCOLADAEAMYQNV 958
Db	907 TCWDADPLKRPSTSKQIVOLIEKQISDSTNHIYSNL 941
RESULT 14	
Q98SUI	
ID	Q98SUI PRELIMINARY; PRT; 977 AA.
AC	Q98SUI;
DT	01-JUN-2001 (TrEMBLrel. 17, Created)
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE	Receptor tyrosine kinase Fms.
OS	Danio nigrofasciatus.
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC	Cyprinidae; Danio.
OX	NCBI_TaxID=144739;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=21359118; PubMed=11466528;
RA	Parichy D.M., Johnson S.L.;
RT	"zebrafish hybrids suggest genetic mechanisms for pigment pattern
RT	diversification in Danio.";
RL	Dev. Genes Evol. 211:319-328(2001).
CC	-!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC	-!- SIMILARITY: BELONGS TO THE CSP-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC	PROTEIN KINASES.
DR	EMBL; AF324481; RAK15303.1; --
DR	HSSP; P11362; IFGK.
DR	InterPro; IPR007110; IG-like.
DR	InterPro; IPR003598; IG_c2.
DR	InterPro; IPR003006; IG_MHC.
DR	InterPro; IPR000719; Prot_kinase.
DR	InterPro; IPR001824; RTKinaseIII.
DR	InterPro; IPR001245; Tyr_kinase.
DR	Pfam; PF00047; ig_4.
DR	Pfam; PF00069; pkinase; 1.
DR	ProDom; PD000001; Prot_kinase; 2.
DR	SMART; SM00408; IGC2; 1.
DR	SMART; SM00219; TyrKc; 1.
DR	PROSITE; PS50835; IG LIKE; 3.
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR	PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR	PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW	ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW	Phosphorylation; Receptor; Transferrase; Transmembrane;
KW	Tyrosine-protein kinase.
FT	VARIANT 60 60 V -> L.
FT	VARIANT 103 103 A -> S.
FT	VARIANT 742 742 P -> S.
FT	VARIANT 899 899 E -> D.
SQ	SEQUENCE 977 AA; 109881 MW; CF54E129FAB10E3E CRC64;
Query Match 23.3%; Score 1230.5; DB 13; Length 977;	
Best Local Similarity 32.8%; Pred. No. 9.9e-95;	
Matches 326; Conservative 174; Mismatches 367; Indels 127; Gaps 34;	
QY	71 SSGTYVEAAAEVDVSASITLQVLVDAPGNISCL-WVPKHSLLNCQPHFDLQNRGVSMV 129
Db	27 NSGAL---AGTDVLDDSGSLQVCEGDPVTFPRVAKH-----KRYTSKRVGKIRSPR 78
QY	130 ILKMTETQAGRY-LLFIQSEATNYTILFTVSIRNT-LLY-----TLRRPYFRKQENODAL 182
Db	79 VEKATVDFTGYKYCYVINGSNLSASSVHVFRDSEVLFPVSPSSLR--YVRK-EGEDLL 135
QY	183 V-CISESVPEP-IVEWVLCDQSGESCKEESPAWVKKEKVL-----HELFGTDIRCARN 236
Db	136 LPCL---LTDPTDPTDFRMDNGSAAPYGMNATPDRKGLIRNVHPGFNADYICARIG 192
QY	237 LGRECTRLFTIDLQNTPTTLQPLK-----VGEPLWIRCKAVHNVHGFGLTWELN 289
Db	193 GAERKVSKEIPFNVITQRLRFP-PYVYLKENEYVKLVGERLQISCTTNPNFYNNVTWTHSS 251
QY	290 KALBEGNYFEMSTYSTNRTMIRILPAFVSVARNDTGYTCTSSSKHPSQSALVT---IVG 346
Db	252 KRLPKPE--EKSTWEGDRLAIESILT--ISSVQLSDTGNITCTGQNEAGANSSTTQLLVVD 308
QY	347 KGFI-----NATNSSEDEYIDQYEFCEFCFSVRFKAYFQIRC-TW---TFSRKSFPFCBQ 394



```

QY 598 EYFVDPREYDYDKWEPRENLEFGKVLGSAFGKVMNATAYGISKTGVSIOQAVKMLK 647
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
563 NYVYIDPTQLFYDHKWEFPRNRLSFGKTLGAGAFKGVVEATAYGLIKSDAAMTAVKMLK 622
QY 648 EKADSSREALMSLKMTQLGSHENIVNLLGACTLSQPIYLIFBYCCYGDLLNVLRSKR 707
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
623 PSAHLTREALMSLKVLSYLGNHMNIIVNLLGACTVGGPTLVITEYCCYGDLLNVLRSKR 682
QY 708 EKFRHTWTE-----IPKEHNFSFYPTFQSHPNSSM---PGSREVOIHPDSDOISGLHGN 758
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
683 DSFICSKQEDHAEALYKNLLHSHKSSCNDSTNEYMDMKPGVSV-VPTKTKRRAARIG 741
QY 759 SFHSEDEIYENQKLEEEEDLNVLTPEDLCPAYQVAKGMEFLFKSCVHRDLAARNVL 818
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
742 SY-----IERDVAPSIMEDDEL-ALDLEDLSFSYQVAKGWAFKASKNCIHRDLAARNIL 795
QY 819 VTHCKVVKICDPGLARDIMSDSNVYVVGNAFLPVKMWAPESLPEGIYTIKSDVMSYGILL 878
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
796 LTHGRITKICDFGLARDIKNDSNVYVVGNAFLPVKMWAPESIPNCVYTFESDVMSYGIFL 855
QY 879 WEIFSLGWNYPGIPVDANFYKLIQNGFKMDQPFYATEEIIYIMQSCWAPDSRKPSPFN 938
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
856 WELFSLGSSYPGMPVDSKFKYKMKEGFRMLSPHAPAEYDINKTCWDADPLKRPTFKQ 915
QY 939 LTSFLGCLADAEAEAMYNQVNDGRVSECPHTYQN 971
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
916 IVOLIEKQISDTNQIYSN----LANCSPROEN 944

```

Search completed: August 26, 2003, 07:27:52  
Job time : 87.5421 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2003, 07:13:34 ; Search time 62.5315 Seconds  
(without alignments)  
2520.578 Million cell updates/sec

Title: US-09-919-408A-4  
Perfect score: 5274  
Sequence: 1 MPALARDAGTVPLLVFSAM.....PFSREMDLGLSPQAQVEDS 993

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: A\_Geneseq\_19Jun03.\*  
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
25: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5274	100.0	993	16	AA67816 Flk2 receptor prot
2	5274	100.0	993	16	AA67536 Human flk-2. Homo
3	5274	100.0	993	17	AA67419 Murine foetal live
4	5274	100.0	993	18	AAW19873 Human flk-2 recept
5	5274	100.0	993	20	AAW08617 Human flk-2 protei
6	5274	100.0	993	23	ABG70916 Human receptor pro
7	5274	100.0	993	23	AAE25819 Human receptor pro
8	5271	99.9	993	14	AA67503 Human flk-2. Homo
9	5266	99.8	993	14	AA67495 Human flk-2 recept

10	5266	99.8	993	16	AA675961 Human STX-1. Homo
11	5265	99.8	993	16	AA67816 Human flk2/fit3 ty
12	5262	99.8	1167	14	AA67376 Human flk-2. Homo
13	5159	97.8	983	19	AAW63588 Human receptor typ
14	5157.5	97.8	986	19	AAW63587 Human receptor typ
15	5157.5	97.8	986	19	AAW63589 Human receptor typ
16	5153.5	97.7	994	19	AAW63586 Human receptor typ
17	4533.5	86.0	1000	16	AA67816 Flk2/fit3 tyrosine
18	4429.5	84.0	992	14	AA67502 Murine flk-2. Mus
19	4429.5	84.0	992	16	AA67815 Flk2 receptor prot
20	4429.5	84.0	992	16	AA67535 Mouse flk-2. Mus
21	4429.5	84.0	992	17	AA67418 Human foetal liver
22	4429.5	84.0	992	18	AAW19874 Murine flk-2 recep
23	4429.5	84.0	992	20	AAW08616 Murine flk-2 prote
24	4429.5	84.0	992	23	ABG70915 Mouse receptor pro
25	4429.5	84.0	992	23	AAE25818 Murine receptor pr
26	4421.5	83.8	992	14	AA67494 Murine flk-2 recep
27	4421.5	83.8	992	14	AA67375 Murine flk-2. Mus
28	4404.5	83.5	992	13	AA678038 Human flk-2. Mus
29	3323	63.0	665	19	AA67585 Human receptor typ
30	2747	52.1	749	19	AAW78002 Protein pMON32390
31	1640	31.1	481	15	AA67579 Soluble Flk-2. Mu
32	1578	29.9	307	23	ABB81198 Human FLT3 protein
33	1286	24.4	977	21	AAV51322 Bovine c-Kit bk-1
34	1266	24.0	975	22	AAE07144 Murine Kit/stem ce
35	1266	24.0	975	22	AAE07148 Mutant murine Kit/
36	1251	23.7	976	22	AAE07145 Human Kit/stem cel
37	1251	23.7	976	22	AAE07149 Mutant human Kit/s
38	1251	23.7	976	22	AAU00375 Human stem cell gr
39	1251	23.7	976	24	ABP72761 Human c-kit. Homo
40	1251	23.7	976	24	ABR39481 Human c-kit wild-t
41	1251	23.7	976	24	ABP72341 Human c-kit. Homo
42	1251	23.7	976	24	ABP72342 Human c-kit. Homo
43	1251	23.7	976	24	ABP57411 Human c-kit. Protei
44	1251	23.7	976	24	ABP57418 Human c-kit protei
45	1251	23.7	976	24	ABP96025 Human c-kit protei

## ALIGNMENTS

RESULT 1  
AA67816  
ID AA67816 standard; Protein; 993 AA.  
XX XX  
AC AA67816;  
XX XX  
DT 25-MAR-2003 (updated)  
DT 18-AUG-1995 (first entry)  
XX XX  
DE Flk2 receptor protein-tyrosine-kinase.  
XX XX  
KW Human Flk2; receptor protein-tyrosine-kinase; primitive  
KW hemopoietic cell; fetal liver kinase; diagnostic ligand  
KW isolation; bone marrow disease therapy.  
XX XX  
OS Homo sapiens.  
XX XX  
FH Key Location/Qualifiers  
FT Peptide 1..27 /note= "signal peptide"  
FT Protein 28..993 /note= "mature protein"  
FT Domain 28..543 /note= "extracellular receptor domain"  
FT Domain 544..563 /note= "transmembrane region"  
FT Domain 564..993 /note= "intracellular catalytic domain"  
XX XX  
PD W09500554-A2.  
PD 05-JAN-1995.

XX PF 17-JUN-1994; 94WO-0506944.  
XX PR 18-JUN-1993; 93US-0080244.  
XX PR 21-JUN-1993; 93US-0081508.  
XX PR 23-NOV-1993; 93US-0157490.  
XX (UYPR-) UNIV PRINCETON.  
XX PA Lemischka IR;  
XX PI WPI; 1995-052014/07.  
XX DR N-PSDB; AAQ81013.  
XX  
PT Ligand for receptor protein tyrosine kinase - useful for the  
PT stimulation of primitive haematopoietic stem cells causing  
PT proliferation and/or differentiation  
XX  
XX Disclosure; Fig 1b; 13lpp; English.  
XX  
XX The sequence corresponds to a human Flk2 (fetal liver kinase)  
CC receptor protein-tyrosine-kinase, which is expressed in primitive  
CC hematopoietic cells but not in mature hematopoietic cells. The  
CC protein is useful in isolation of receptor ligands, which have  
CC applications in diagnosis of bone marrow disorders and in  
CC stimulating proliferation and/or differentiation of primitive  
CC hematopoietic stem cells.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX  
XX Sequence 993 AA;  
SQ

Query Match 100.0%; Score 5274; DB 16; Length 993;  
Best Local Similarity 100.0%; Pred. NO. 0;  
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPLLVFVSAMIFGTITNQDLPIVKVLINHKNDSSVYKSSSPMUSEP 60  
DB 1 MPALARDAGTVPLLVFVSAMIFGTITNQDLPIVKVLINHKNDSSVYKSSSPMUSEP 60  
QY 61 EDLGCALRPOSSGTVEAAVVDVSAITLQVLVDAPGNISCLWFKHSLNCQPHFDL 120  
DB 61 EDLGCALRPOSSGTVEAAVVDVSAITLQVLVDAPGNISCLWFKHSLNCQPHFDL 120  
QY 121 QNRGVVSMVILKMTQAGEYLLFIQSEATNYTLFTVSRNTLLYTLRRPYFRKMNQD 180  
DB 121 QNRGVVSMVILKMTQAGEYLLFIQSEATNYTLFTVSRNTLLYTLRRPYFRKMNQD 180  
QY 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240  
DB 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240  
QY 241 CTRLEFDLNOTPOTTLPOLFLKVGEPFLWTRCKAVHNVHGFGLTWELNKALEBGNYPFEM 300  
DB 241 CTRLEFDLNOTPOTTLPOLFLKVGEPFLWTRCKAVHNVHGFGLTWELNKALEBGNYPFEM 300  
QY 301 STYSTNRMTIRILFAFVSSVARNDTGYTCCSSRHPQSQALVTIVGKGFINATNSSDYE 360  
DB 301 STYSTNRMTIRILFAFVSSVARNDTGYTCCSSRHPQSQALVTIVGKGFINATNSSDYE 360  
QY 361 IDQYEEFCFSVRKAYPQIRCTWTFRSKSPCEQKGLDNGYSISKPCNHKHQPGEYIFHA 420  
DB 361 IDQYEEFCFSVRKAYPQIRCTWTFRSKSPCEQKGLDNGYSISKPCNHKHQPGEYIFHA 420  
QY 421 ENDDAQTKMFTLIRRKPVLAASQASQCFSDGYPSPWTKKCDSPNCTEITE 480  
DB 421 ENDDAQTKMFTLIRRKPVLAASQASQCFSDGYPSPWTKKCDSPNCTEITE 480  
QY 481 GVNRRKANRKFVGOWSSSTILNMEAIGFLVKCCAYNSLGTSCETILLNSPGFPFIQD 540  
DB 481 GVNRRKANRKFVGOWSSSTILNMEAIGFLVKCCAYNSLGTSCETILLNSPGFPFIQD 540  
QY 541 NISFYATIGVCLLFIVVLTLLCHIKYKQFRYESQLQMVQVTGSSDNEYFYVDFREYD 600

Db 541 NISFYATIGVCLLFIVVLTLLCHIKYKQFRYESQLQMVQVTGSSDNEYFYVDFREYD 600  
QY 601 LKWEFPRENLEFGKVLGSGAFKVMNATAYGISTQVAVTMLKEKADSSEREALMS 660  
Db 601 LKWEFPRENLEFGKVLGSGAFKVMNATAYGISTQVAVTMLKEKADSSEREALMS 660  
QY 661 ELKMMTOLGSHENIVNLLGACTLSGPIYLIFFEYCCYGDLLNLYRSKREKPHRTWTEIFKE 720  
Db 661 ELKMMTOLGSHENIVNLLGACTLSGPIYLIFFEYCCYGDLLNLYRSKREKPHRTWTEIFKE 720  
QY 721 HNFSPYPTQSHPNSSMPGSRVQIHPDSQISGLHNSFHSDEIEYENQKLEEBEDL 780  
Db 721 HNFSPYPTQSHPNSSMPGSRVQIHPDSQISGLHNSFHSDEIEYENQKLEEBEDL 780  
QY 781 NVLTPEDLICFAYQVAKGMFELEFKSCVHRDLAARNVLTGHKVKVICDFGLARDINSDS 840  
Db 781 NVLTPEDLICFAYQVAKGMFELEFKSCVHRDLAARNVLTGHKVKVICDFGLARDINSDS 840  
QY 841 NYVVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIFSLGVNYPYGPVPDANFYK 900  
Db 841 NYVVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIFSLGVNYPYGPVPDANFYK 900  
QY 901 LIQNGFKMDQPPYATEEIIIMOSWAFDSRKRPSFPNLTSLFLGCOLADAEAMYQNVDG 960  
Db 901 LIQNGFKMDQPPYATEEIIIMOSWAFDSRKRPSFPNLTSLFLGCOLADAEAMYQNVDG 960  
QY 961 RVSECPHTYQNRPPFSREMDLGLLSPQAQVEDS 993  
Db 961 RVSECPHTYQNRPPFSREMDLGLLSPQAQVEDS 993

RESULT 2  
AAR67536  
ID AAR67536 standard; Protein; 993 AA.  
XX AAR67536;  
XX 25-MAR-2003 (updated)  
DT 04-JUL-1995 (first entry)  
XX Human flk-2.  
XX Fetal liver kinase-2; flk-2; protein tyrosine-kinase receptor;  
KW hematopoiesis; stem cell.  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FT Peptide 1..27 /label= Sig\_peptide  
FT Domain 28..543 /label= Extracellular\_receptor\_domain  
FT Region 544..563 /label= Transmembrane\_region  
FT Domain 564..993 /label= Intracellular\_catalytic\_domain  
XX US5367057-A.  
XX 22-NOV-1994.  
XX 30-APR-1993; 93US-0055269.  
XX 02-APR-1991; 91US-0679666.  
XX 28-JUN-1991; 91US-0728913.  
XX 15-NOV-1991; 91US-0793065.  
XX 24-DEC-1991; 91US-0813593.  
XX 26-JUN-1992; 92US-0906397.  
XX 12-NOV-1992; 92US-0975049.  
XX 19-NOV-1992; 92US-0977451.  
XX 30-APR-1993; 93US-0055269.  
XX (UYPR-) UNIV PRINCETON.  
XX PA

XX PI Lemischka IR;  
 XX WPI; 1995-005894/01.  
 DR N-PSDB; AAQ79069.  
 XX Murine flk-2 receptor protein tyrosine kinase - used to stimulate  
 PT proliferation and/or stimulation of primitive mammalian  
 PT haematopoietic stem cells in vitro or in vivo.  
 XX Disclosure; Fig. 2A-1F; 69pp; English.  
 XX cDNAs encoding receptor protein tyrosine-kinases, mouse fetal liver  
 CC kinase-2 (flk-2), human flk-2 and mouse flk-1 are given in AAQ79068-70,  
 CC respectively, and the deduced amino acid sequences in AAQ67535-37,  
 CC respectively.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX Sequence 993 AA;  
 SQ

Query Match 100.0%; Score 5274; DB 16; Length 993;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPLLVFVSAMIFGTITNODLPVTKVCLINHKNDSSVYKSSSYPMVSESP 60  
 DB 1 MPALARDAGTVPLLVFVSAMIFGTITNODLPVTKVCLINHKNDSSVYKSSSYPMVSESP 60  
 QY 61 EDGCGALRPOSSGTVEAAAVEVDVSASITLOVLVDAPGNISCLWVFKHSNLCQPHFDL 120  
 DB 61 EDGCGALRPOSSGTVEAAAVEVDVSASITLOVLVDAPGNISCLWVFKHSNLCQPHFDL 120  
 QY 121 QNRGVSWILKMTQAGEYLLFIQSEATNTYTLFTVSINTLLYTLRPPYFRKMNQD 180  
 DB 121 QNRGVSWILKMTQAGEYLLFIQSEATNTYTLFTVSINTLLYTLRPPYFRKMNQD 180  
 QY 181 ALVCISSEVPEPVEVWVLCDSQGESCKEESPAVVKKEKVLHFGTDIRCCARNEILGRE 240  
 DB 181 ALVCISSEVPEPVEVWVLCDSQGESCKEESPAVVKKEKVLHFGTDIRCCARNEILGRE 240  
 QY 241 CTLRFTIDLNOTPOTTLQPLFKVGEPLWIRCKAVHNVHGFLTWELNKALEBGNFEM 300  
 DB 241 CTLRFTIDLNOTPOTTLQPLFKVGEPLWIRCKAVHNVHGFLTWELNKALEBGNFEM 300  
 QY 301 STYSTNRTMIRILFAFVSSVARNDTGYTTCSSSKHPSQSALVTIVGKGFNATNSSDYE 360  
 DB 301 STYSTNRTMIRILFAFVSSVARNDTGYTTCSSSKHPSQSALVTIVGKGFNATNSSDYE 360  
 QY 361 IDQYEEFCFVRKAYPOIRCTWTFRSKSPCEQKGLONGYSISKPNHKHQGEYIFHA 420  
 DB 361 IDQYEEFCFVRKAYPOIRCTWTFRSKSPCEQKGLONGYSISKPNHKHQGEYIFHA 420  
 QY 421 ENDDAQTKMFTLNIRKPOVLAEASASQSCFSDGYPLPSWTWKKCDKSPNCTEITE 480  
 DB 421 ENDDAQTKMFTLNIRKPOVLAEASASQSCFSDGYPLPSWTWKKCDKSPNCTEITE 480  
 QY 481 GWNRKANRVFGOWVSSSTLNMSEAIGPLVKCAVNSLGTSETILLNSPGPFPIQD 540  
 DB 481 GWNRKANRVFGOWVSSSTLNMSEAIGPLVKCAVNSLGTSETILLNSPGPFPIQD 540  
 QY 541 NISFYATIGVCLLPIVLTLLIHKYKQFRYESQLQWQVQTGSSDNEYFYVDFREYD 600  
 DB 541 NISFYATIGVCLLPIVLTLLIHKYKQFRYESQLQWQVQTGSSDNEYFYVDFREYD 600  
 QY 601 LKWEFFRENLEFGKVLGSGAGFKVMNATAYGISKTGYSIQAVKMLKEKADSSREALMS 660  
 DB 601 LKWEFFRENLEFGKVLGSGAGFKVMNATAYGISKTGYSIQAVKMLKEKADSSREALMS 660  
 QY 661 ELKMTQOLGSHENTVNLGACTLSGPIYLIPEYCCYGDLLNLYLRSKREKPHRTWTEIPKE 720  
 DB 661 ELKMTQOLGSHENTVNLGACTLSGPIYLIPEYCCYGDLLNLYLRSKREKPHRTWTEIPKE 720  
 QY 721 HNFSFYPTFQSHPNSSMPGSRVQIHPDSQIISGLHGNSPHSEDEIYENQKRLSEEDL 780

DB 721 HNFSFYPTFQSHPNSSMPGSRVQIHPDSQIISGLHGNSPHSEDEIYENQKRLSEEDL 780  
 QY 781 NVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKIKCDFGLARDIMSDS 840  
 DB 781 NVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKIKCDFGLARDIMSDS 840  
 QY 841 NYVVRGNARLPVVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNPPYGPDPVANDFYK 900  
 DB 841 NYVVRGNARLPVVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNPPYGPDPVANDFYK 900  
 QY 901 LIQNGFKMDQPPYATEEIIYIMQSCWAFDSKRPSPFNLTSLFLGCQLADAEEMVQNVYD 960  
 DB 901 LIQNGFKMDQPPYATEEIIYIMQSCWAFDSKRPSPFNLTSLFLGCQLADAEEMVQNVYD 960  
 QY 961 RVSECPHTYQNRPPFRREMOLGLLSPOAQVEDS 993  
 DB 961 RVSECPHTYQNRPPFRREMOLGLLSPOAQVEDS 993

RESULT 3  
 AAR97419  
 ID AAR97419 standard; Protein; 993 AA.  
 XX AC AAR97419;  
 XX 25-MAR-2003 (updated)  
 DT 11-DEC-1996 (first entry)  
 XX Murine foetal liver kinase 2.  
 XX Murine; foetal liver kinase 2; flk-2; protein tyrosine kinase;  
 KW monoclonal; antibody; extracellular domain; receptor assay;  
 KW haematopoietic stem cell; ligand; stimulation; proliferation;  
 KW differentiation; treatment; anaemia; bone marrow damage;  
 KW cancer chemotherapy; radiation.  
 XX Mus musculus.  
 OS  
 XX Key  
 FH Peptide 1..27 Location/Qualifiers  
 FT Peptide /label= sig\_peptide  
 FT Peptide 28..993  
 FT Peptide /label= mat\_peptide  
 FT Domain 28..544 /label= extracellular\_domain  
 FT Domain 545..564 /label= transmembrane\_domain  
 FT Domain 565..993 /label= intracellular\_domain  
 PN US548065-A.  
 XX 20-AUG-1996.  
 XX 31-OCT-1994; 94US-0252517.  
 XX 19-NOV-1992; 92US-0977451.  
 XX 02-APR-1991; 91US-0679666.  
 XX 28-JUN-1991; 91US-0728913.  
 XX 15-NOV-1991; 91US-0793065.  
 XX 24-DEC-1991; 91US-0813593.  
 XX 26-JUN-1992; 92US-0906397.  
 XX 12-NOV-1992; 92US-0975049.  
 XX 30-APR-1993; 93US-0055269.  
 XX 31-OCT-1994; 94US-0252517.  
 XX (UYPR-) UNIV PRINCETON.  
 XX Lemischka IR;  
 XX WPI; 1996-392678/39.  
 DR N-PSDB; AAT38734.  
 DR



XX Anti-fetal liver kinase 2 (flk-2) antibodies - useful in assays,  
PT for isolating haematopoietic stem cells expressing receptor and for  
PT obtaining ligands  
XX  
PS Claim 1; Columns 39-48; 50pp; English.  
XX  
CC The present sequence is murine foetal liver kinase 2 (flk-2),  
CC a protein tyrosine kinase. Isolated antibodies, pref. monoclonal,  
CC raised against the extracellular portion of flk-2 can be used to  
CC assay for flk receptors on the surface of primitive haematopoietic  
CC stem cells, and to isolate positive cells. The antibodies can also  
CC be used as, or to obtain ligands, which stimulate the proliferation  
CC and/or differentiation of stem cells. The ligands can be used, e.g.  
CC for treating anaemia, or bone marrow damage resulting from cancer  
CC chemotherapy, or radiation.  
CC (Updated on 25-MAR-2003 to correct PF field.)  
XX  
SQ Sequence 993 AA;

Query Match 100.0%; Score 5274; DB 17; Length 993;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPALARDAGTVPLLVFSGAMIFGTTNQDLPVVKVNLNHNKNDSSVGKSSYPMVSESP 60  
Db 1 MPALARDAGTVPLLVFSGAMIFGTTNQDLPVVKVNLNHNKNDSSVGKSSYPMVSESP 60  
Qy 61 EDLGCALPQSSGTYEAAAEVDVSASITLQVLVDAPGNISCLVWFKHSSLNCPHFDL 120  
Db 61 EDLGCALPQSSGTYEAAAEVDVSASITLQVLVDAPGNISCLVWFKHSSLNCPHFDL 120  
Qy 121 QNRGVSVMLKMTQAGEYLLFQSEATNTYILFTVSRINTLYLRRPFRQWENQD 180  
Db 121 QNRGVSVMLKMTQAGEYLLFQSEATNTYILFTVSRINTLYLRRPFRQWENQD 180  
Qy 181 ALVCISESVPEPIVEMVLCDGSGCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240  
Db 181 ALVCISESVPEPIVEMVLCDGSGCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240  
Qy 241 CTRFTIDNLQTPQTLFLKVGCEPLWIRCKAVVNHGFLTWELNKALEEGNYFEM 300  
Db 241 CTRFTIDNLQTPQTLFLKVGCEPLWIRCKAVVNHGFLTWELNKALEEGNYFEM 300  
Qy 301 STYSTNRMTWIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFINATNSSEDEY 360  
Db 301 STYSTNRMTWIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFINATNSSEDEY 360  
Qy 361 IDQVEEFCFVRFKAYPOIRCTWTFPSRKSFPCEQKGLDNGYSISKFCNHKHQPGYIFHA 420  
Db 361 IDQVEEFCFVRFKAYPOIRCTWTFPSRKSFPCEQKGLDNGYSISKFCNHKHQPGYIFHA 420  
Qy 421 ENDDAQFTMTFLNTRRRPOVLAESAQSCFSDGYPLPSWTWKCKDKNCTEEITE 480  
Db 421 ENDDAQFTMTFLNTRRRPOVLAESAQSCFSDGYPLPSWTWKCKDKNCTEEITE 480  
Qy 481 GWNKRNKRVFGQWSSVSTLNMBSAIGFLVKCAVNSLGTSCETILLNSPGPPFFIQD 540  
Db 481 GWNKRNKRVFGQWSSVSTLNMBSAIGFLVKCAVNSLGTSCETILLNSPGPPFFIQD 540  
Qy 541 NISFYATIGVCLLFIVLTLCHYKXKQFRIESQVMQVQTGSSDNEYFYVDFPREYED 600  
Db 541 NISFYATIGVCLLFIVLTLCHYKXKQFRIESQVMQVQTGSSDNEYFYVDFPREYED 600  
Qy 601 LKWEPPRENLEPGKVLGSGAFQKVNATAYGISTGVSIQVAVKMLKEKADSSERREALMS 660  
Db 601 LKWEPPRENLEPGKVLGSGAFQKVNATAYGISTGVSIQVAVKMLKEKADSSERREALMS 660  
Qy 661 ELKMMTQLGSHENIVNLGACTLSGPYILIFVCCYGLDNLNLRKREKFRHTWTWEIFKE 720  
Db 661 ELKMMTQLGSHENIVNLGACTLSGPYILIFVCCYGLDNLNLRKREKFRHTWTWEIFKE 720  
Qy 721 HNFSPYPTFQSHPNSSMPGSRVQIHPSDDQISGLHGNSFHSEDEIETENQKRLBEEDL 780

Db 721 HNFSPYPTFQSHPNSSMPGSRVQIHPSDDQISGLHGNSFHSEDEIETENQKRLBEEDL 780  
Qy 781 NVLTFEDLLCFAYQVAKGMEFEFKSCVHRDLAARNVLVTHGKVKIKICDFGLARDIMSDS 840  
Db 781 NVLTFEDLLCFAYQVAKGMEFEFKSCVHRDLAARNVLVTHGKVKIKICDFGLARDIMSDS 840  
Qy 841 NYVVRGNARLPVKWMAPESLFEGYITIKSDVWSYGILLWEIFSLGVNYPGIPVDANFYK 900  
Db 841 NYVVRGNARLPVKWMAPESLFEGYITIKSDVWSYGILLWEIFSLGVNYPGIPVDANFYK 900  
Qy 901 LIQNGFKMDQPFYATEEYIIIMQSCWAFDSRKRSPFNLTSLFLGCQLADAEAMTQNVDS 960  
Db 901 LIQNGFKMDQPFYATEEYIIIMQSCWAFDSRKRSPFNLTSLFLGCQLADAEAMTQNVDS 960  
Qy 961 RVSECPHTYQNRPRPFSREMDLGLLSPQAQVEDS 993  
Db 961 RVSECPHTYQNRPRPFSREMDLGLLSPQAQVEDS 993

RESULT 4  
AAW19873  
ID AAW19873 standard; Protein; 993 AA.  
XX  
AC AAW19873;  
DT 25-MAR-2003 (updated)  
DT 19-AUG-1997 (first entry)  
XX  
DE Human flk-2 receptor.  
XX  
KW Human; fetal liver kinase 2; flk2; receptor protein tyrosine kinase;  
KW ptk; liver; spleen; thymus; adult; brain; bone marrow; primitive portion;  
KW haematopoietic hierarchy; extracellular domain; soluble form; ligand;  
KW proliferation; differentiation; mammalian; haematopoietic stem cell;  
KW macrocytic anaemia; aplastic anaemia; cancer; radiation; mouse.  
XX  
OS Homo sapiens.  
FH Key  
FT Peptide 1..27  
FT Protein /note= "Signal peptide"  
FT /note= "Mature flk-2"  
XX US5621090-A.  
XX 15-APR-1997.  
XX 26-JUN-1992; 92US-0906397.  
XX 26-JUN-1992; 92US-0906397.  
XX 02-APR-1991; 91US-0679666.  
XX 28-JUN-1991; 91US-0728913.  
XX 15-NOV-1991; 91US-0793065.  
XX 24-DEC-1991; 91US-0813593.  
XX (UYPR-) UNIV PRINCETON.  
XX Lemischka IR;  
XX WPI; 1997-235228/21.  
XX N-PSDB; AAT72117.  
XX  
PT Protein containing the extracellular domain of human flk-2 - used  
PT for identification of primitive haematopoietic cell proliferation  
PT and differentiation stimulatory ligands, e.g. for treating anaemia  
XX  
PS Claim 1; Fig 1B; 55pp; English.  
XX  
CC This sequence represents human fetal liver kinase 2 (flk2). flk-2 is  
CC a receptor protein tyrosine kinase (ptk) and is important in transducing  
CC putative self-renewal signals from the environment. flk-2 is expressed

CC in fetal liver, spleen, fetal thymus, adult brain and adult bone marrow,  
CC and it is thought that flk-2 is expressed in the entire primitive portion  
CC of the haematopoietic hierarchy. The invention concerns a recombinant  
CC nucleic acid, preferably mRNA, which encodes a protein containing only  
CC the extracellular domain of human flk-2 and lacking the flk-2 intra-  
CC cellular catalytic domain. The resultant protein represents a soluble  
CC form of flk-2 which is used to isolate specific ligands for flk-2. These  
CC ligands can be used to stimulate proliferation and/or differentiation of  
CC mammalian haematopoietic stem cells, in vivo or in vitro, e.g. for  
CC treatment of macrocytic or aplastic anaemia or bone marrow damage caused  
CC by cancer treatment or radiation.  
CC (Updated on 25-MAR-2003 to correct PF field.)  
XX Sequence 993 AA;  
SQ Query Match 100.0%; Score 5274; DB 18; Length 993;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPALARDAGTVPLLVVFSAMIFGTITNQDLPVVKVLIHKNNDSSVGKSSYPMVSESP 60  
Db 1 MPALARDAGTVPLLVVFSAMIFGTITNQDLPVVKVLIHKNNDSSVGKSSYPMVSESP 60

Qy 61 EDIGCALRPOSSGTVYAAAVEVDVSAITLOVLVDAPGNISCLWVKHSSLNCPHFDL 120  
Db 61 EDIGCALRPOSSGTVYAAAVEVDVSAITLOVLVDAPGNISCLWVKHSSLNCPHFDL 120

Qy 121 QNRGVSMVLKMTQAGEVLLFQSEATNTYTLFTVSIIRNTLLYLRFPYFRKMNQD 180  
Db 121 QNRGVSMVLKMTQAGEVLLFQSEATNTYTLFTVSIIRNTLLYLRFPYFRKMNQD 180

Qy 181 ALVCISESVPEPIVEWVLCDSQGSCKEESPAVVKKEKVLHFLFGTDIRCARNELGRE 240  
Db 181 ALVCISESVPEPIVEWVLCDSQGSCKEESPAVVKKEKVLHFLFGTDIRCARNELGRE 240

Qy 241 CTRFTTDLNQTPTTLPOLFLKVGEPWIRCKAVHNVHNGFGLTWELNKALEGNYFEM 300  
Db 241 CTRFTTDLNQTPTTLPOLFLKVGEPWIRCKAVHNVHNGFGLTWELNKALEGNYFEM 300

Qy 301 STYSTNRTMIRILFAVSSVARNDTGYTTCSSSKHPSQALVTIVGKGFINATNSSEDE 360  
Db 301 STYSTNRTMIRILFAVSSVARNDTGYTTCSSSKHPSQALVTIVGKGFINATNSSEDE 360

Qy 361 IDQVEEFCFVRFRKAYPOIRCTWTFSRKSFCEQKGLDNGYSISKFCNHKHPQGEYIFPA 420  
Db 361 IDQVEEFCFVRFRKAYPOIRCTWTFSRKSFCEQKGLDNGYSISKFCNHKHPQGEYIFPA 420

Qy 421 ENDDAQFTKMTLIRRKPOVLAEASASQSCFSDGYPLPSWTWKCSKSPNCTEEITE 480  
Db 421 ENDDAQFTKMTLIRRKPOVLAEASASQSCFSDGYPLPSWTWKCSKSPNCTEEITE 480

Qy 481 GWNRKANRKYFGQWSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPPFFIQD 540  
Db 481 GWNRKANRKYFGQWSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPPFFIQD 540

Qy 541 NISFVATTIGVCLLFTWLTLICHYKKQFRYESQLOQWQVVTGSSDNEYFYVDPREYED 600  
Db 541 NISFVATTIGVCLLFTWLTLICHYKKQFRYESQLOQWQVVTGSSDNEYFYVDPREYED 600

Qy 601 LKWEPPRENLEFGKVLGSGAFGKVNATAYISKTGVSIOQAVKMLKADSSEREALMS 660  
Db 601 LKWEPPRENLEFGKVLGSGAFGKVNATAYISKTGVSIOQAVKMLKADSSEREALMS 660

Qy 661 ELKMTQLGSHENIVNLGACTLSGPVILIFPYCCYGDLNLYLRKREKFRHTWTETPKE 720  
Db 661 ELKMTQLGSHENIVNLGACTLSGPVILIFPYCCYGDLNLYLRKREKFRHTWTETPKE 720

Qy 721 HNFSPYPTFQSHPNSSMPGSRVQIHPDSQISGLHGNFSHSEDEIYENQRLSEEDL 780  
Db 721 HNFSPYPTFQSHPNSSMPGSRVQIHPDSQISGLHGNFSHSEDEIYENQRLSEEDL 780

Qy 781 NVLTPEDLCLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDS 840

Db 781 NVLTPEDLCLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDS 840

Qy 841 NYVVRGNARLPVKWMAPESLPEGIYTIKSDVWSYGILLWEIFSLGVNYPFGIPVDANFYK 900  
Db 841 NYVVRGNARLPVKWMAPESLPEGIYTIKSDVWSYGILLWEIFSLGVNYPFGIPVDANFYK 900

Qy 901 LIQNGFKMDQPPFYATBEIYIIMQSWAFDSRKRPSFPNLTSLFLGCQLADAEAEAMYQNVGD 960  
Db 901 LIQNGFKMDQPPFYATBEIYIIMQSWAFDSRKRPSFPNLTSLFLGCQLADAEAEAMYQNVGD 960

Qy 961 RVSECPHTYQNRPPFSREMDLGLLSPOAQVEDS 993  
Db 961 RVSECPHTYQNRPPFSREMDLGLLSPOAQVEDS 993

RESULT 5  
AAAY08617  
ID AAAY08617 standard; Protein; 993 AA.  
XX AC AAAY08617;  
XX DT 05-AUG-1999 (first entry)  
XX DE Human flk-2 protein.  
XX KW Human; flk-2; flk-1; cell isolation; fetal liver kinase; receptor;  
XX KM monoclonal; polyclonal; antibody; tyrosine kinase.  
XX OS Homo sapiens.  
XX PN US912133-A.  
XX PD 15-JUN-1999.  
XX PF 10-FEB-1998; 98US-0021324.  
XX PR 19-NOV-1992; 92US-0977451.  
XX PR 02-APR-1991; 91US-0679686.  
XX PR 28-JUN-1991; 91US-0728913.  
XX PR 15-NOV-1991; 91US-0793065.  
XX PR 24-DEC-1991; 91US-0813593.  
XX PR 26-JUN-1992; 92US-0906397.  
XX PR 12-NOV-1992; 92US-0975049.  
XX PR 30-APR-1993; 93US-0055289.  
XX PR 31-OCT-1994; 94US-0252498.  
XX PR 15-FEB-1996; 96US-0601891.  
XX (UYPR-) UNIV PRINCETON.  
XX Lemischka IR;  
XX WPI; 1999-357194/30.  
XX N-PSDB; AAX77515.  
XX Isolating hematopoietic cells expressing fetal liver kinase 1  
PT receptors  
XX Disclosure; Pig 1b; 59pp; English.  
XX This invention describes a novel method of isolating cells expressing  
CC fetal liver kinase 1 (flk-1) receptors on their surface and comprises  
CC binding the cells to a polyclonal or monoclonal antibody specific to the  
CC flk-1 receptor and isolating the cells that have bound to the  
CC antibody. The method can be used to isolate hematopoietic stem cells in  
CC any mammal but preferably a rat, mouse, rabbit or human. The proteins of  
CC the invention belong to the receptor protein family. This sequence  
CC represents the human flk-2 protein which is used in the method of the  
CC invention.  
XX Sequence 993 AA;  
SQ Query Match 100.0%; Score 5274; DB 20; Length 993;  
Best Local Similarity 100.0%; Pred. No. 0;

```
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MPALARDAGTVLLVVFSAIPIGTITNQDLPIVKVLINHKNDSSVGKSSYPMVSESP 60
Db 1 MPALARDAGTVLLVVFSAIPIGTITNQDLPIVKVLINHKNDSSVGKSSYPMVSESP 60
Qy 61 EDLGCALRPOSSGTVEAAAEVDVSASTTLOVLVDAPGNISCLVWFKHSSLNCPHFDL 120
Db 61 EDLGCALRPOSSGTVEAAAEVDVSASTTLOVLVDAPGNISCLVWFKHSSLNCPHFDL 120
Qy 121 QNRGVSVVILKWTQAGEYLLFTQSEATNTILFTVSIRNTLYTLRRPFRKMNQD 180
Db 121 QNRGVSVVILKWTQAGEYLLFTQSEATNTILFTVSIRNTLYTLRRPFRKMNQD 180
Qy 181 ALVCISEVPPIVWVLCDSGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240
Db 181 ALVCISEVPPIVWVLCDSGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240
Qy 241 CTRLEFTIDLNTQPTTLQFLKVGEPWLIRCKAVVHNGFGLTWELNKALEEGNYFEM 300
Db 241 CTRLEFTIDLNTQPTTLQFLKVGEPWLIRCKAVVHNGFGLTWELNKALEEGNYFEM 300
Qy 301 STYSTNRTWIRILPAFVSSVARNDGYTTCSSSKHPSQSALVTIIVKGFINATNSSEDE 360
Db 301 STYSTNRTWIRILPAFVSSVARNDGYTTCSSSKHPSQSALVTIIVKGFINATNSSEDE 360
Qy 361 IDQYEEFCFSVRFKAYPOIRCTWTTSRKSFPCEQKGLDNGYSISKFCNHKHPGGEYI 420
Db 361 IDQYEEFCFSVRFKAYPOIRCTWTTSRKSFPCEQKGLDNGYSISKFCNHKHPGGEYI 420
Qy 421 ENDDAQFTKMTLNTRRPQVLAESAQSCFSDGYPLPSWTWKCKSDKSPNCTEITE 480
Db 421 ENDDAQFTKMTLNTRRPQVLAESAQSCFSDGYPLPSWTWKCKSDKSPNCTEITE 480
Qy 481 GVNRKANRKFVGQWVSSSTLMSAIGFLVKCCAYNSLGTSCETIILNSGPPFPFD 540
Db 481 GVNRKANRKFVGQWVSSSTLMSAIGFLVKCCAYNSLGTSCETIILNSGPPFPFD 540
Qy 541 NISFYATIGVCLLFTIVLTLTICHYKKQFRYESQLQWQVTSQSSDNEYFYVDREYD 600
Db 541 NISFYATIGVCLLFTIVLTLTICHYKKQFRYESQLQWQVTSQSSDNEYFYVDREYD 600
Qy 601 LKWEPPRENLEFGKVLGSGAGFKVNNATAYGISKTGVSIOAVKMLKKAQSSEREA 660
Db 601 LKWEPPRENLEFGKVLGSGAGFKVNNATAYGISKTGVSIOAVKMLKKAQSSEREA 660
Qy 661 ELKMTQLGSHENIVNLGACTLSGPVLIPEYCCYGDLLNVLRSKRKEFHTWTW 720
Db 661 ELKMTQLGSHENIVNLGACTLSGPVLIPEYCCYGDLLNVLRSKRKEFHTWTW 720
Qy 721 HNFSPYPTFQSHPNSSMPGSRVQIHPDSQDQISGLHGNFSHSEDEIYENQKRLB 780
Db 721 HNFSPYPTFQSHPNSSMPGSRVQIHPDSQDQISGLHGNFSHSEDEIYENQKRLB 780
Qy 781 NVLTFEDLLCPAYQAKGMEFLPKSCVHRDLAARNVLVTHGKVKICDGLARDIMSD 840
Db 781 NVLTFEDLLCPAYQAKGMEFLPKSCVHRDLAARNVLVTHGKVKICDGLARDIMSD 840
Qy 841 NVVVRGNARLPVKWAPESLPEGIYTIKSDVWSYGILLWEIPSLGWNYPGIPVDAN 900
Db 841 NVVVRGNARLPVKWAPESLPEGIYTIKSDVWSYGILLWEIPSLGWNYPGIPVDAN 900
Qy 901 LIQNGFKMDQFPYATEEYIIMQSCWAFDSRKRPSFPNLTSLGCOLADAEAMQYND 960
Db 901 LIQNGFKMDQFPYATEEYIIMQSCWAFDSRKRPSFPNLTSLGCOLADAEAMQYND 960
Qy 961 RVSECPHTYQNRPPSRMDLGLLSPQAQVEDS 993
Db 961 RVSECPHTYQNRPPSRMDLGLLSPQAQVEDS 993
```

RESULT 6  
ABG70916

```
ID XX ABG70916 standard; Protein; 993 AA.
AC XX ABG70916;
DT XX 09-DEC-2002 (first entry)
DE XX Human receptor protein tyrosine kinase, FLK-2.
KW Human; FLK-2; foetal liver kinase; antianaemic; enzyme;
KW ophthalmological; receptor protein tyrosine kinase; aplastic anaemia;
KW primitive haematopoietic cell; stem cell; macrocytic anaemia;
KW bone marrow damage; cancer chemotherapy.
OS XX Homo sapiens.
FH XX Key Location/Qualifiers
FT Peptide 1..27
FT Protein /label= Signal_peptide
FT 28..993
FT /label= Mature_FLK_2
PN XX US2002119545-A1.
PD XX 29-AUG-2002.
PF XX 01-JUN-2001; 2001US-0872136.
PR XX 19-NOV-1992; 92US-0977451.
PR XX 30-APR-1993; 93US-0055269.
PR XX 31-OCT-1994; 94US-0252498.
PR XX 15-FEB-1996; 96US-0601891.
PR XX 10-FEB-1998; 98US-0021324.
PR XX 10-DEC-1998; 98US-0208786.
PR XX 02-APR-1991; 91US-0679666.
PR XX 28-JUN-1991; 91US-0728913.
PR XX 15-NOV-1991; 91US-0793065.
PR XX 24-DEC-1991; 91US-0813593.
PA XX (LEMI/) LEMISCHKA I R.
PI XX Lemischka IR;
XX WPI; 2002-731356/79.
DR N-PSDB; ABS55045.
PT New protein tyrosine kinase expressed in primitive hematopoietic cells
PT (HC) and not expressed in mature HC, and ligands for the protein,
PT useful for stimulating proliferation of primitive hematopoietic stem
PT cells -
XX Claim 37; Fig 1b; 64pp; English.
CC The invention relates to a receptor protein tyrosine kinase (pTK)
CC expressed in primitive haematopoietic cells and not expressed in mature
CC haematopoietic cells, and named FLK-1 and -2 (foetal liver kinase).
CC Also included are the nucleic acids encoding the FLK proteins, FLK
CC expression vectors, a ligand that binds to human or murine FLK-2 or FLK-1
CC and stimulates the proliferation and/or differentiation of the primitive
CC haematopoietic cells and a murine cell line 2018 having American type
CC Culture Collection (ATCC) accession number ATCC CRL 10907. The ligands
CC are useful for stimulating the proliferation and/or differentiation of
CC primitive mammalian haematopoietic stem cells. The receptor pTK molecules
CC are useful for stimulating the self-renewal of the totipotent
CC haematopoietic stem cell and to stimulate the development of all cells of
CC the haematopoietic system both in vitro and in vivo. The ligands for the
CC receptors act as haematopoietic growth factors. The ligands are useful in
CC treating humans whose primitive stem cells do not sufficiently undergo
CC self-renewal e.g. macrocytic and aplastic anaemia and bone marrow damage
CC resulting from cancer chemotherapy and radiation. The present
CC sequence represents human FLK-2.
XX Sequence 993 AA;
```

Query Match	100.0%;	Score 5274;	DB 23;	Length 993;
Best Local Similarity	100.0%;	Prod. No. 0;		
Matches 993;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1	MPALARDAGTVPLLVAFVSAMIFGTTNQDLPVIKCVLINHKNDSSVGKSSSYPMWSESP	60
Db	1	MPALARDAGTVPLLVAFVSAMIFGTTNQDLPVIKCVLINHKNDSSVGKSSSYPMWSESP	60
Qy	61	EDLGCALRPOSSGTVYEAFAAVEVDVSASITLQVLVDAPGNISCLVWFVKHSSLNCQHPFDL	120
Db	61	EDLGCALRPOSSGTVYEAFAAVEVDVSASITLQVLVDAPGNISCLVWFVKHSSLNCQHPFDL	120
Qy	121	QNRGVSWVLKMTETQAGEVLLFIQSEATNTYTLFTVSIINTLLYTLRRPYFRKMENQD	180
Db	121	QNRGVSWVLKMTETQAGEVLLFIQSEATNTYTLFTVSIINTLLYTLRRPYFRKMENQD	180
Qy	181	ALVCISSESVPEPIVEWVLCSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNEILGRE	240
Db	181	ALVCISSESVPEPIVEWVLCSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNEILGRE	240
Qy	241	CTRFLTIDLNQTPQTLPLQFLKVGEPILWRCKAVHVNHGFLGTWELNKALEEGNYFEM	300
Db	241	CTRFLTIDLNQTPQTLPLQFLKVGEPILWRCKAVHVNHGFLGTWELNKALEEGNYFEM	300
Qy	301	STYSTNRTMIRILFAFVSSVARNDTGYTCSSSHXPQSALVTIVKGFINATNSSSEDE	360
Db	301	STYSTNRTMIRILFAFVSSVARNDTGYTCSSSHXPQSALVTIVKGFINATNSSSEDE	360
Qy	361	IDQYEEFCFSVRPKAYPOIRCTWTFSRKSFCQEQKGLDNGYSISKFCNHHKQPGCEYIFPHA	420
Db	361	IDQYEEFCFSVRPKAYPOIRCTWTFSRKSFCQEQKGLDNGYSISKFCNHHKQPGCEYIFPHA	420
Qy	421	ENDDAQFTKMPFTLNIRRKPOVLASASQAQSCFSDGYPPLPSWTWKKCKSDKSPNCTEEITE	480
Db	421	ENDDAQFTKMPFTLNIRRKPOVLASASQAQSCFSDGYPPLPSWTWKKCKSDKSPNCTEEITE	480
Qy	481	GVNWRKANRKVFQGWVSSSTLNMSEAIGKFLVKCCAYNSLGTSCETILLNSPGPPPIQD	540
Db	481	GVNWRKANRKVFQGWVSSSTLNMSEAIGKFLVKCCAYNSLGTSCETILLNSPGPPPIQD	540
Qy	541	NISFYATIGVCLLFIIVLTLLIICHYKKQPYESQOLQMVQVTGSSDNRYFYVDREYTD	600
Db	541	NISFYATIGVCLLFIIVLTLLIICHYKKQPYESQOLQMVQVTGSSDNRYFYVDREYTD	600
Qy	601	LKWEPFRENLEFGKVLGSGAFQKWNATAYGISKTGVSIOAVKMLKEKADSSSREALMS	660
Db	601	LKWEPFRENLEFGKVLGSGAFQKWNATAYGISKTGVSIOAVKMLKEKADSSSREALMS	660
Qy	661	ELKQMTQLGSHENIVNLGACTLSGPYILIFEYCCYGDLNLYLRSKREKFKHRTWTEIFKE	720
Db	661	ELKQMTQLGSHENIVNLGACTLSGPYILIFEYCCYGDLNLYLRSKREKFKHRTWTEIFKE	720
Qy	721	HNFSFYPTFQSHPNSSMPGSRVQIHPDSQDISGLHGNFSHSEDEIYEYENQKLEEEEDL	780
Db	721	HNFSFYPTFQSHPNSSMPGSRVQIHPDSQDISGLHGNFSHSEDEIYEYENQKLEEEEDL	780
Qy	781	NVLTFEDLLCPAYQAKGMEFLFKSCVHRDLAARNVLTHTGKVVKICDFGLARDIMSDS	840
Db	781	NVLTFEDLLCPAYQAKGMEFLFKSCVHRDLAARNVLTHTGKVVKICDFGLARDIMSDS	840
Qy	841	NYVVRGNARLPVKWMAPESLPEGIYTIKSDVWSYGILLWEIFSLGWNYPGIPVDANFYK	900
Db	841	NYVVRGNARLPVKWMAPESLPEGIYTIKSDVWSYGILLWEIFSLGWNYPGIPVDANFYK	900
Qy	901	LIQNGFKMDQPFYATEEIIYIMQSCWAFDSRRKPSFPNLTSPGLGQLADEAEMAYQNVDG	960
Db	901	LIQNGFKMDQPFYATEEIIYIMQSCWAFDSRRKPSFPNLTSPGLGQLADEAEMAYQNVDG	960
Qy	961	RVSECPHTYQNRPPFSRMDLGLLSPOAQVEDS	993
Db	961	RVSECPHTYQNRPPFSRMDLGLLSPOAQVEDS	993

RESULT 7	
AAE25819	AAE25819 standard; Protein; 993 AA.
XX AC	
XX AC	AAE25819;
XX	
DT	15-NOV-2002 (first entry)
XX	
DE	Human receptor protein tyrosine kinase, flk-2.
XX	
KW	Human; receptor protein tyrosine kinase; pTK; haema
KW	growth factor; therapeutic; macrocytic anaemia; apl
KW	bone marrow damage; cancer; chemotherapy; radiation
KW	flk-2 protein.
XX	
OS	Homo sapiens.
XX	
FH	Location/Qualifiers
FT	1..27
FT	/label= Signal-peptide
FT	28..993
FT	/note= "Human mature flk-2 protein"
FT	1..516
FT	/note= "Extracellular domain"
FT	517..536
FT	/note= "Transmembrane domain"
FT	537..966
FT	/note= "Intracellular domain"
XX	
PN	US2002072077-A1.
XX	
PD	13-JUN-2002.
XX	
PF	31-JUL-2001; 2001US-0919408.
XX	
PR	19-NOV-1992; 92US-0977451.
PR	30-APR-1993; 93US-0035269.
PR	31-OCT-1994; 94US-0252498.
PR	15-FEB-1996; 96US-0601891.
PR	10-FEB-1998; 98US-0021324.
PR	10-DEC-1998; 98US-0208786.
PR	02-APR-1991; 91US-0679666.
PR	28-JUN-1991; 91US-0728913.
PR	15-NOV-1991; 91US-0793065.
PR	24-DEC-1991; 91US-0813593.
XX	
PA	(LEWI/) LEMISCHKA I R.
XX	
PI	Lemischka IR;
XX	
DR	WPI; 2002-607237/65.
XX	
DR	N-PSDB; AAD42484.
XX	
PT	New protein tyrosine kinase expressed in primitive
PT	(HC) and not expressed in mature HC, and ligands for
PT	stimulating proliferation of primitive hematopoietic
XX	
PS	Claim 37; Page 25-28; 64pp; English.
XX	
CC	The present invention relates to receptor protein t
CC	expressed in primitive haematopoietic cells and nob
CC	haematopoietic cells, polynucleotides encoding such
CC	for the protein. Ligands which bind to pTK of the i
CC	for stimulating the proliferation and/or differentia
CC	mammalian haematopoietic stem cells. The receptor p
CC	useful for stimulating the self renewal of the totip
CC	stem cell and to stimulate the development of all c
CC	system both in vitro and in vivo. The ligands for c
CC	haematopoietic growth factors. The ability of the l
CC	proliferation of stem cells both in vitro and in vi
CC	therapeutic applications such as treating humans w
CC	cells do not sufficiently undergo self-renewal. It
CC	conditions that occur when defects in haematopoiet
CC	

CC related growth factors depress the number of white blood cells such as macrocytic and aplastic anaemia or bone marrow damage resulting from cancer chemotherapy and radiation. The present sequence is human receptor pTK, flk-2.									
XX	Sequence	993 AA;							
QY	Query Match	100.0%;	Score 5274;	DB 23;	Length 993;				
CC	Best Local Similarity	100.0%;	Pred. No. 0;						
CC	Matches 993;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	MPALARDAGTVPVLLVVFSA	MTGTTT	QDLPV	KVCLNHKNDSSV	GKSSSYPMVSESP	60		
DB	1	MPALARDAGTVPVLLVVFSA	MTGTTT	QDLPV	KVCLNHKNDSSV	GKSSSYPMVSESP	60		
QY	61	EDLGCALRPQSSGTYEAAAEV	DVSASIT	TLQVL	VDAPGNISCLW	VFKHSSLNCPHFDL	120		
DB	61	EDLGCALRPQSSGTYEAAAEV	DVSASIT	TLQVL	VDAPGNISCLW	VFKHSSLNCPHFDL	120		
QY	121	QNRGVSMVILKMTETQAGEY	LLFTQSE	ATNTY	TLFTVSR	INTLYTLRRPYFRKMNQD	180		
DB	121	QNRGVSMVILKMTETQAGEY	LLFTQSE	ATNTY	TLFTVSR	INTLYTLRRPYFRKMNQD	180		
QY	181	ALVCISESVPEPIVEMVIL	CDSGESCKE	SPAVK	KEEVLHEL	FGTDIRCCARNELGRE	240		
DB	181	ALVCISESVPEPIVEMVIL	CDSGESCKE	SPAVK	KEEVLHEL	FGTDIRCCARNELGRE	240		
QY	241	CTRFTIDNLTPOPTTLQ	FLKVGCEPLW	IRKCAV	VHNGFGLT	WELENKALEEGNYEM	300		
DB	241	CTRFTIDNLTPOPTTLQ	FLKVGCEPLW	IRKCAV	VHNGFGLT	WELENKALEEGNYEM	300		
QY	301	STYSTNRTMIRILFAFV	SSVARNDT	GYTCS	SSKHPQSAL	VTIIVKGFINATNSEDYE	360		
DB	301	STYSTNRTMIRILFAFV	SSVARNDT	GYTCS	SSKHPQSAL	VTIIVKGFINATNSEDYE	360		
QY	361	IDQYBEFCFSVRFKAY	POIRCTWT	SRKSP	CEQKGLDNG	YSISKFCNHKHPGQYI	420		
DB	361	IDQYBEFCFSVRFKAY	POIRCTWT	SRKSP	CEQKGLDNG	YSISKFCNHKHPGQYI	420		
QY	421	ENDDAQFTKMTFLN	TRRPQVLA	ASASQ	ASCFSDGYP	PLPSMTWKCDKSPNCTE	480		
DB	421	ENDDAQFTKMTFLN	TRRPQVLA	ASASQ	ASCFSDGYP	PLPSMTWKCDKSPNCTE	480		
QY	481	GYWNRKANRKFVGQ	WSSSTLNM	SEAIK	GFVKCCAY	NSLGTSCETILLNSPG	540		
DB	481	GYWNRKANRKFVGQ	WSSSTLNM	SEAIK	GFVKCCAY	NSLGTSCETILLNSPG	540		
QY	541	NISFYATIGVCLLFI	VWLTLLI	CHYK	KQPRYESQ	LQWQVGTGSSDNEYFY	600		
DB	541	NISFYATIGVCLLFI	VWLTLLI	CHYK	KQPRYESQ	LQWQVGTGSSDNEYFY	600		
QY	601	LKWEFPRENLEFG	KVLGSAFG	KVMNAT	AYGISKT	GVSTQVAVKMLKEKAD	660		
DB	601	LKWEFPRENLEFG	KVLGSAFG	KVMNAT	AYGISKT	GVSTQVAVKMLKEKAD	660		
QY	661	ELKMTQLGSHENI	VNLLGACT	LSGP	IYLI	FEYCCYGLDNLNLR	720		
DB	661	ELKMTQLGSHENI	VNLLGACT	LSGP	IYLI	FEYCCYGLDNLNLR	720		
QY	721	HNFSYPTFQSHPN	SMPSRE	VQIHP	DSQISGLH	GNFSHSEDEIENQRL	780		
DB	721	HNFSYPTFQSHPN	SMPSRE	VQIHP	DSQISGLH	GNFSHSEDEIENQRL	780		
QY	781	NVLTREDLLCFAY	QVAKGM	FELEF	SKVHRD	LAARNVLTHGKV	840		
DB	781	NVLTREDLLCFAY	QVAKGM	FELEF	SKVHRD	LAARNVLTHGKV	840		
QY	841	NYWVRGNARLP	VKNWAPES	LFEG	ITYIKSD	VMSYGILLWEIF	900		
DB	841	NYWVRGNARLP	VKNWAPES	LFEG	ITYIKSD	VMSYGILLWEIF	900		
QY	901	LIQNGFKMDQPPY	ATEEIIY	IMQSC	WAFDSR	KRPSFNLTS	960		

Db	901	LIQNGFKMDQPPY	ATEEIIY	IMQSC	WAFDSR	KRPSFNLTS	960		
Qy	961	RVSECPHTYQNR	PPFSREMD	LGLLSPQ	AVQVEDS	993			
Db	961	RVSECPHTYQNR	PPFSREMD	LGLLSPQ	AVQVEDS	993			
RESULT 8									
AAAR37503									
ID	AAAR37503	standard;	Protein;	993	AA.				
XX	AAAR37503;								
XX	25-MAR-2003	(updated)							
DT	19-OCT-1993	(first entry)							
XX	Human flk-2.								
XX	Murine; receptor;	protein;	tyrosine kinase;	ptk;	flk-2;	primitive;			
KW	hematopoietic cell;	mature;	family;	conserved;	region;				
KW	catalytic domain;	c-kit;	fetal liver kinase;	flk;	fetal;	spleen;			
KW	thymus; adult;	brain;	bone marrow;	multipotential;	CFU-Blast colony;				
XX	hierarchy;	transduction;	T-lymphoid;	lineage.					
OS	Homo sapiens.								
Key	Location/Qualifiers								
FT	1..27								
FT	/note= "Hydrophobic leader sequence"								
FT	28...993								
FT	/note= "Mature murine flk-2"								
FT	28...543								
FT	/note= "Extracellular receptor domain"								
FT	544..563								
FT	/note= "Transmembrane region"								
FT	564..993								
FT	/note= "Intracellular catalytic domain"								
PN	WO9310136-A1.								
XX	27-MAY-1993.								
XX	16-NOV-1992;	92WO-US09893.							
XX	15-NOV-1991;	91US-0793065.							
XX	(UYPR-) UNIV PRINCETON.								
XX	Lemischka IR;								
XX	WPI; 1993-182479/22.								
XX	N-PSDB; AAQ40915.								
XX	Totipotent haematopoietic stem cell receptors, their ligands and DNA sequences - for treating anaemia(s) and bone marrow damage due to e.g. cancer chemotherapy or radiotherapy								
PS	Claim 39; Fig 1b; 127pp; English.								
CC	This sequence represents the human receptor protein tyrosine kinase (ptk), flk-2. The nucleic acid encoding this receptor is expressed in primitive hematopoietic cells and not in mature hematopoietic cells. Members of this family of ptk's can be recognised by the conserved amino acid regions in the catalytic domain. This family of ptk's also contains c-kit. These new receptors are termed fetal liver kinases (flk's) after the tissue in which they were discovered. flk-2 is also expressed in fetal spleen, fetal thymus, adult brain and adult bone marrow. flk-2 is expressed in individual multipotential CFU-Blast colonies capable of generating numerous multilineage colonies upon replating. It is likely therefore that flk-2 is expressed in the entire primitive portion of the hematopoietic hierarchy. This is consistent with flk-2 being important in transducing putative self-renewal signals from the environment. flk-2 is the first receptor								

CC pTK known to be expressed in the T-lymphoid lineage.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 993 AA;

Query Match		99.9%; Score 5271; DB 14; Length 993;
Best Local Similarity		99.9%; Pred. No. 0;
Matches 992; Conservative 1; Mismatches 0; Indels 0; Gaps 0;		
QY	1	MPALARDAGTVPLLVAFVSAMIFGTTQDLPVIRKVLINHKNDSSVGKSSYPWVSESP 60
DB	1	MPALARDAGTVPLLVAFVSAMIFGTTQDLPVIRKVLINHKNDSSVGKSSYPWVSESP 60
QY	61	EDLGCALRPOSSGTVYAAAEEVDVSASITLQVLVDAPGNISCLWPKHSSLNCPHFDL 120
DB	61	EDLGCALRPOSSGTVYAAAEEVDVSASITLQVLVDAPGNISCLWPKHSSLNCPHFDL 120
QY	121	QNRGVSWILKMTETQAGEYLLFIQSEATNYTLFTVSIRNTLLYLRPRYPRMENQD 180
DB	121	QNRGVSWILKMTETQAGEYLLFIQSEATNYTLFTVSIRNTLLYLRPRYPRMENQD 180
QY	181	ALVCISESVPEPIVEWVLCDQSGESCKEESPVAVYKKEKVLHFLGTDIRCCARNELGRE 240
DB	181	ALVCISESVPEPIVEWVLCDQSGESCKEESPVAVYKKEKVLHFLGTDIRCCARNELGRE 240
QY	241	CTRLFTIDLNOTPOTTLPLQFLKVGEPILWIRKAVHVNHGFGLTWLENKALBEGNYFEM 300
DB	241	CTRLFTIDLNOTPOTTLPLQFLKVGEPILWIRKAVHVNHGFGLTWLENKALBEGNYFEM 300
QY	301	STYSTNRTRMIRILFAFVSSVARNDTGYTCTSSSKHPSQSALVTIVGKGFINATNSSBDE 360
DB	301	STYSTNRTRMIRILFAFVSSVARNDTGYTCTSSSKHPSQSALVTIVGKGFINATNSSBDE 360
QY	361	IDQYEFCEFSVRKAPQIQRCTWTFPSRSPCEQKGLDNGYSISKFNKHQPGCEYIFHA 420
DB	361	IDQYEFCEFSVRKAPQIQRCTWTFPSRSPCEQKGLDNGYSISKFNKHQPGCEYIFHA 420
QY	421	ENDDAQFTKMTINIRKPOVLAEASASQASQSDGYPLPSWTWKCKSDKSPNCTEITE 480
DB	421	ENDDAQFTKMTINIRKPOVLAEASASQASQSDGYPLPSWTWKCKSDKSPNCTEITE 480
QY	481	GVNWRKANRKFQGWVSSSTLNSEAIKGFVKCCAYNSLGTSCETILLNSPGFPFPIQD 540
DB	481	GVNWRKANRKFQGWVSSSTLNSEAIKGFVKCCAYNSLGTSCETILLNSPGFPFPIQD 540
QY	541	NISFYATIGVCLLFIIVVLTLLI CHYKKQPRYESQLQMVQVTSDDNEYFYVDFREYED 600
DB	541	NISFYATIGVCLLFIIVVLTLLI CHYKKQPRYESQLQMVQVTSDDNEYFYVDFREYED 600
QY	601	LKWEPPRENLEFGKVLGSGAFGKVMNATAYISKTGVSIOAVKMLKEKADSSEREAALMS 660
DB	601	LKWEPPRENLEFGKVLGSGAFGKVMNATAYISKTGVSIOAVKMLKEKADSSEREAALMS 660
QY	661	ELKQMTQLGSHENIVNLLGACTLSGPIYLLIFCYCCVGDLLNLYRSKREKPHRTWTIFKE 720
DB	661	ELKQMTQLGSHENIVNLLGACTLSGPIYLLIFCYCCVGDLLNLYRSKREKPHRTWTIFKE 720
QY	721	HNFSFYPTQSHPNSSMPGSRVQIHPDSQISGLHNSFHSDEIEYENQKLEEEEDL 780
DB	721	HNFSFYPTQSHPNSSMPGSRVQIHPDSQISGLHNSFHSDEIEYENQKLEEEEDL 780
QY	781	NVLTPEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHKVKICDPGLARDIMSDS 840
DB	781	NVLTPEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHKVKICDPGLARDIMSDS 840
QY	841	NVTVRGNARLPVKWMAPELIFEGYITIKSDVWSYGILLWEIFSLGVPYPGIIVDANFYK 900
DB	841	NVTVRGNARLPVKWMAPELIFEGYITIKSDVWSYGILLWEIFSLGVPYPGIIVDANFYK 900
QY	901	LQNGFKMDQPPFYATEIYIIMQSCWAFDSRKRKPSPNLTSLFLGCQLADAEAMYNQVNDG 960
DB	901	LQNGFKMDQPPFYATEIYIIMQSCWAFDSRKRKPSPNLTSLFLGCQLADAEAMYNQVNDG 960

QY	961	RVSECPHTYQNRPPFSREMDLGLLSPOAQVEDS 993
DB	961	RVSECPHTYQNRPPFSREMDLGLLSPOAQVEDS 993
RESULT 9		
AAR44995		
ID	AAR44995	standard; Protein; 993 AA.
XX	AAR44995;	
AC	AAR44995;	
XX	25-MAR-2003	(updated)
DT	27-JUN-1994	(first entry)
XX	Human flk-2 receptor protein tyrosine kinase.	
XX	Receptor protein tyrosine kinase; pTK family; foetal liver kinase; hflk; primitive; totipotent; haematopoietic cell; stem cell; proliferation; stromal cell.	
KW	Homo sapiens.	
OS	Homo sapiens.	
XX	Key	Location/Qualifiers
PH	Peptide	1..27
FT		/label= signal sequence
FT		/note= "hydrophobic leader"
FT	Protein	28..993
FT		/label= flk-2
FT		/note= "mature protein"
FT	Domain	28..544
FT		/label= extracellular_domain
FT	Region	545..563
FT		/label= transmembrane_region
FT	Domain	564..993
FT		/label= intracellular_catalytic_domain
XX	US5270450-A.	
XX	14-DEC-1993.	
XX	19-NOV-1992;	92US-0977451.
XX	02-APR-1991;	91US-0679666.
PR	28-JUN-1991;	91US-0728913.
PR	15-NOV-1991;	91US-0793065.
PR	24-DEC-1991;	91US-0813593.
PR	26-JUN-1992;	92US-0906397.
PR	12-NOV-1992;	92US-0975049.
PR	19-NOV-1992;	92US-0977451.
XX	(UYPR-) UNIV PRINCETON.	
PA	Lemischka IR;	
PI	WPI; 1993-405021/50.	
DR	N-PSDB; AA053503.	
XX	Isolated nucleic acid molecules of hematopoietic stem cell receptor flk-2 - encoding mammalian receptor protein tyrosine kinases expressed in primitive haematopoietic cells	
XX	Disclosure; Fig 1b; 60pp; English.	
XX	Nucleic acid sequences coding for murine flk-2 and specified subfragments of it are claimed. The human flk-2 coding sequence (i.e. AAQ53503) is also disclosed. The murine and human flk-2 polypeptides represent a new class of receptor protein tyrosine kinases which are expressed only in primitive haematopoietic cells. (Updated on 25-MAR-2003 to correct PF field.)	
XX	Sequence	993 AA;
SQ	Query Match	99.8%; Score 5266; DB 14; Length 993;

Best Local Similarity 99.9%; Pred. No. 0;			
Matches 992; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
QY	1	MPALARDAGTVP	LLVVFSA
Db	1	MPALARDAGTVP	LLVVFSA
QY	61	EDLGCALRPQSSG	TVYEA
Db	61	EDLGCALRPQSSG	TVYEA
QY	121	QNRGVVSMVILK	MTETQAGEYLL
Db	121	QNRGVVSMVILK	MTETQAGEYLL
QY	181	ALVCISES	VEPIE
Db	181	ALVCISES	VEPIE
QY	241	CTRLFTID	LNQPTTLP
Db	241	CTRLFTID	LNQPTTLP
QY	301	STYSTNRT	WIRILFA
Db	301	STYSTNRT	WIRILFA
QY	361	IQOYEEFC	SVRFKAYP
Db	361	IQOYEEFC	SVRFKAYP
QY	421	ENDDAQ	FTKMTINIRK
Db	421	ENDDAQ	FTKMTINIRK
QY	481	GWNRKANR	KVFQWYSS
Db	481	GWNRKANR	KVFQWYSS
QY	541	NTSFYAT	IGVCLLF
Db	541	NTSFYAT	IGVCLLF
QY	601	LKWEFFRE	NLEFGK
Db	601	LKWEFFRE	NLEFGK
QY	661	ELKMMTQ	LGSHE
Db	661	ELKMMTQ	LGSHE
QY	721	HNFSFYPT	FQSHNS
Db	721	HNFSFYPT	FQSHNS
QY	781	NVLTFED	LLCFAYQ
Db	781	NVLTFED	LLCFAYQ
QY	841	NYVVRGNAR	LPUKMA
Db	841	NYVVRGNAR	LPUKMA
QY	901	LIONGFMD	OPFYAT
Db	901	LIONGFMD	OPFYAT
QY	961	RVSECPHT	YQNR
Db	961	RVSECPHT	YQNR

AAR75961			
ID	AAR75961 standard; Protein; 993 AA.		
XX			
AC	AAR75961;		
XX			
DT	25-MAR-2003 (updated)		
DT	29-DEC-1995 (first entry)		
XX			
DE	Human STK-1.		
XX			
KW	STK-1; receptor PTK; protein tyrosine kinase.		
XX			
OS	Homo sapiens.		
XX			
FH	Key		
FT	Peptide		
FT	/label= signal		
FT	/note= "unique PTK receptor sequence"		
FT	25..26		
FT	Cleavage-site		
FT	Domain		
FT	35..524		
FT	/note= "Part of extracellular domain contg. 22 Cys residues marked on Fig. 1"		
FT	Modified-site		
FT	43..45		
FT	/label= potential N-linked glycosylation site		
FT	Modified-site		
FT	100..102		
FT	/label= see above		
FT	Modified-site		
FT	151..153		
FT	/label= see above		
FT	Modified-site		
FT	250..252		
FT	/label= see above		
FT	Modified-site		
FT	306..308		
FT	/label= see above		
FT	Modified-site		
FT	323..325		
FT	/label= see above		
FT	Modified-site		
FT	351..353		
FT	/label= see above		
FT	Modified-site		
FT	354..356		
FT	/label= see above		
FT	Modified-site		
FT	473..475		
FT	/label= see above		
FT	Modified-site		
FT	502..504		
FT	/label= see above		
FT	Modified-site		
FT	542..562		
FT	/label= transmembrane spanning region		
FT	Region		
FT	708..782		
FT	/label= kinase insert region		
FT	Domain		
FT	617..622		
FT	/label= ATP binding domain		
FT	Domain		
FT	835..840		
FT	/label= WWAPES motifs		
FT	/note= "cytoplasmic domain"		
FT	Peptide		
FT	808..813		
FT	/note= "used to design PCR oligos"		
FT	Peptide		
FT	870..875		
FT	/note= "used to design PCR oligos"		
XX	W09519175-A1.		
PN			
XX	20-JUL-1995.		
PD			
XX			
PF	06-JAN-1995; 95WO-US00176.		
XX			
PR	14-JAN-1994; 94US-0183211.		
XX			
PA	(UJJO ) UNIV JOHNS HOPKINS.		
PA	(UJPE-) UNIV PENNSYLVANIA.		
XX			
PI	Civin CI, Gewirtz AM, Small D;		
XX			
DR	WPI; 1995-263709/34.		
DR	N-PSDB; AAQ91356.		
XX			
PT	Artificial STK-1 gene and gene-specific anti:sense oligo:nucleotide		





QY 1 MPALARDAGTVLLVVFSGAMIFGTTINQDLPIKVLINHKNDSSVGKSSYPWVSSESP 60  
 Db 1 MPALARDAGTVLLVVFSGAMIFGTTINQDLPIKVLINHKNDSSVGKSSYPWVSSESP 60  
 QY 61 EDLGCALRPOSSGTYEAAAEVDVDSASTLOVLDAPGNISCLVWFKHSSLNCOHPDPL 120  
 Db 61 EDLGCALRPOSSGTYEAAAEVDVDSASTLOVLDAPGNISCLVWFKHSSLNCOHPDPL 120  
 QY 121 QNRGVSVVILKWTETQAGEYLLFQSEATNTYTLFTVSRNTLYTLRPPFRKXENQD 180  
 Db 121 QNRGVSVVILKWTETQAGEYLLFQSEATNTYTLFTVSRNTLYTLRPPFRKXENQD 180  
 QY 181 ALVCISESVPEPIVENVLCDSCGESCSEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240  
 Db 181 ALVCISESVPEPIVENVLCDSCGESCSEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240  
 QY 241 CTRLEFTIDLNTQPTTLPOLFLKVGEPILWIRKAVHVNHGFLTWELNKALEEGNYFEM 300  
 Db 241 CTRLEFTIDLNTQPTTLPOLFLKVGEPILWIRKAVHVNHGFLTWELNKALEEGNYFEM 300  
 QY 301 STYSTNRMTIRILFAPVSSVARNDTGYTTCSSSKHPSQSALVTIVGKGFINATNSSEDEY 360  
 Db 301 STYSTNRMTIRILFAPVSSVARNDTGYTTCSSSKHPSQSALVTIVGKGFINATNSSEDEY 360  
 QY 361 IDQVEEFCFSVRFKAYPQIRCTWTFSRKSPCEQKGLDNGYSISKFCNKHQPGGEYIPHA 420  
 Db 361 IDQVEEFCFSVRFKAYPQIRCTWTFSRKSPCEQKGLDNGYSISKFCNKHQPGGEYIPHA 420  
 QY 421 ENDDAQFTQMTNIRRKQVLAESAASQASCFSDGYPLPSWTWKKCSKSPNCTEEITE 480  
 Db 421 ENDDAQFTQMTNIRRKQVLAESAASQASCFSDGYPLPSWTWKKCSKSPNCTEEITE 480  
 QY 481 GYWNRKANRKFQWSSSTLANSEAIKGLFKVCCAYNSLGTSCETILLNSPGPPFIOD 540  
 Db 481 GYWNRKANRKFQWSSSTLANSEAIKGLFKVCCAYNSLGTSCETILLNSPGPPFIOD 540  
 QY 541 NISFVATIGVCLLFIIVLTLILCHYKKQPRYBSQOLQVQVGTSSDNEYFYVDFREYEYD 600  
 Db 541 NISFVATIGVCLLFIIVLTLILCHYKKQPRYBSQOLQVQVGTSSDNEYFYVDFREYEYD 600  
 QY 601 LKWEPPRENLEFGKVLGSGAFGKVMNATYIGISKTGVSIOVAVKMLKEKADSSERREALMS 660  
 Db 601 LKWEPPRENLEFGKVLGSGAFGKVMNATYIGISKTGVSIOVAVKMLKEKADSSERREALMS 660  
 QY 661 ELKMTOLGSHENIVNLGACTLSGPYLIIFYCCYGDLLNLYRSKREKPHRTWTIEPKE 720  
 Db 661 ELKMTOLGSHENIVNLGACTLSGPYLIIFYCCYGDLLNLYRSKREKPHRTWTIEPKE 720  
 QY 721 HNPSFYPTQSHNSMPSGSRVQIHPDSDQISGLHGNSFHSDEIEYENQKLEEBEEDL 780  
 Db 721 HNPSFYPTQSHNSMPSGSRVQIHPDSDQISGLHGNSFHSDEIEYENQKLEEBEEDL 780  
 QY 781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDPGLARDIMSDS 840  
 Db 781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDPGLARDIMSDS 840  
 QY 841 NYVVRGNARLPVKWMAPESEFEGYITKSDVMSYGILLWEIFSLGVNPPYPGIPVDANFYK 900  
 Db 841 NYVVRGNARLPVKWMAPESEFEGYITKSDVMSYGILLWEIFSLGVNPPYPGIPVDANFYK 900  
 QY 901 LIQNGFMKDPFYATEIYIIMOSCAFDSKRPSPFNLTSLFGCOLADAEAEAMYNQVNDG 960  
 Db 901 LIQNGFMKDPFYATEIYIIMOSCAFDSKRPSPFNLTSLFGCOLADAEAEAMYNQVNDG 960  
 QY 961 RVSECPHTYQNRPPFRSREMDLGLLSPQAQVEDS 993  
 Db 961 RVSECPHTYQNRPPFRSREMDLGLLSPQAQVEDS 993

RESULT 12  
 AAR31376  
 ID AAR31376 standard; Protein; 1167 AA.

XX AAR31376;  
 AC 25-MAR-2003 (updated)  
 DT 25-JUN-1993 (first entry)  
 DE Human flk-2.  
 XX Human; receptor; protein; tyrosine kinase; pTK; primitive; mammalian;  
 KW hematopoietic cell; pHC; mature; mHC; fetal; liver kinase 2; flk-2;  
 KW liver; spleen; thymus; adult; brain; marrow; thymocyte; subset;  
 KW multipotential; T-lymphoid; lineage.  
 XX Homo sapiens.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1013 /note= "Nonsense mutation"  
 FT Misc-difference 1038 /note= "Nonsense mutation"  
 FT Misc-difference 1042 /note= "Nonsense mutation"  
 FT Misc-difference 1057 /note= "Nonsense mutation"  
 FT Misc-difference 1102 /note= "Nonsense mutation"  
 FT Misc-difference 1117 /note= "Nonsense mutation"  
 FT Misc-difference 1127 /note= "Nonsense mutation"  
 XX WO9300349-A1.  
 XX 07-JAN-1993.  
 XX 26-JUN-1992; 92WO-US05401.  
 XX 28-JUN-1991; 91US-0728913.  
 PR 15-NOV-1991; 91US-0793065.  
 PR 24-DEC-1991; 91US-0813593.  
 PR 02-APR-1992; 92WO-US02750.  
 XX (UYPR-) UNIV PRINCETON.  
 PA Lemischka IR;  
 PI WPI; 1993-036323/04.  
 DR N-PSDB; AAQ35250.  
 XX Nucleic acid encoding receptor protein tyrosine kinase - allows  
 PT development of ligands to stimulate proliferation and/or  
 PT differentiation of mammalian haematopoietic stem cells  
 XX Claim 9; Fig 1b; 78pp; English.  
 PS This sequence represents a human receptor protein tyrosine kinase  
 CC which belongs to a new functional class of protein tyrosine kinases  
 CC (pTKs). pTKs in this class are expressed in primitive mammalian  
 CC hematopoietic (pHC) cells but not in mature hematopoietic cells (mHC).  
 CC This protein is an example of a receptor pTK and is called fetal liver  
 CC kinase 2 (flk-2). flk-2 is expressed in fetal liver, spleen and thymus,  
 CC most primitive thymocyte subset. Expression of flk-2 mRNA occurs in the  
 CC and adult brain and marrow. Expression of flk-2 is believed to be uncommitted.  
 CC Therefore, thymocytes expressing flk-2 may be multipotential. flk-2  
 CC is the first receptor tyrosine kinase known to be expressed in the  
 CC T-lymphoid lineage.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 1167 AA;

Query Match 99.8%; Score 5262; DB 14; Length 1167;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 991; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPLLVFSAMIFGTITNODLPVVKVLIHKNKNDSSVGKSSSYPMVSESP 60  
 Db 20 MPALARDAGTVPLLVFSAMIFGTITNODLPVVKVLIHKNKNDSSVGKSSSYPMVSESP 79  
 QY 61 EDLGCALRPOSSGTVYEAFAVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCQPHFDL 120  
 Db 80 EDLGCALRPOSSGTVYEAFAVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCQPHFDL 139  
 QY 121 QNRGVSVMLKMTETQAGEYLLFIQSEATNTYILFTVSVIRNTLLYTLRRPYFRKMENQD 180  
 Db 140 QNRGVSVMLKMTETQAGEYLLFIQSEATNTYILFTVSVIRNTLLYTLRRPYFRKMENQD 199  
 QY 181 ALVCISSEVPPIVEWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240  
 Db 200 ALVCISSEVPPIVEWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 259  
 QY 241 CTRLFTIDLNOTPOTTLFQPLFKVGEPLWIRCAVHVNHGFGLTWELNKALBEGNYFEM 300  
 Db 260 CTRLFTIDLNOTPOTTLFQPLFKVGEPLWIRCAVHVNHGFGLTWELNKALBEGNYFEM 319  
 QY 301 STYSTNRTMIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFNATNSSEDEYE 360  
 Db 320 STYSTNRTMIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFNATNSSEDEYE 379  
 QY 361 IDQYEEFCFVRKAYPQIRCTWTFSRKSPPCQKGLDNGYSISKFCNKHQPGYEYIFA 420  
 Db 380 IDQYEEFCFVRKAYPQIRCTWTFSRKSPPCQKGLDNGYSISKFCNKHQPGYEYIFA 439  
 QY 421 ENDDAQFTQMTLNIRRKQVLAELASASQSCFSDGYPLPSWTWKRCSDKSPNCTBEITE 480  
 Db 440 ENDDAQFTQMTLNIRRKQVLAELASASQSCFSDGYPLPSWTWKRCSDKSPNCTBEITE 499  
 QY 481 GWNKRNKRVKGVSSSTLNWSEIKGLFKCCAYNSLGTSCETILLNSPGPFPIQD 540  
 Db 500 GWNKRNKRVKGVSSSTLNWSEIKGLFKCCAYNSLGTSCETILLNSPGPFPIQD 559  
 QY 541 NISFATIGVCLLFIWVLLCHYKQFRYESQLOQVQVGTGSSDNEYFYVDFREYED 600  
 Db 560 NISFATIGVCLLFIWVLLCHYKQFRYESQLOQVQVGTGSSDNEYFYVDFREYED 619  
 QY 601 LKWEPPRENLEFGKVLGSAFGKVMNATAYGISTGVSIGVAVKMLKXKADSSERREALMS 660  
 Db 620 LKWEPPRENLEFGKVLGSAFGKVMNATAYGISTGVSIGVAVKMLKXKADSSERREALMS 679  
 QY 661 ELKMTQLGSHENIYNLLGACTLSGPIYLIIFYCYGDLNLYLRKREKPHRTWEIFKE 720  
 Db 680 ELKMTQLGSHENIYNLLGACTLSGPIYLIIFYCYGDLNLYLRKREKPHRTWEIFKE 739  
 QY 721 HNFSGFYPTFQSHPNSSMPGSRVQIHPDSQISGLHNSFHSEDEIEYENOKRLEBEDL 780  
 Db 740 HNFSGFYPTFQSHPNSSMPGSRVQIHPDSQISGLHNSFHSEDEIEYENOKRLEBEDL 799  
 QY 781 NVLTFEDLLCFAYQVAKMEFLFKSVCHRDLAARNVLVTHGKVKICDFGLARDIMSDS 840  
 Db 800 NVLTFEDLLCFAYQVAKMEFLFKSVCHRDLAARNVLVTHGKVKICDFGLARDIMSDS 859  
 QY 841 NYVVRGNARLPVKWAPSLFEGYITIKSDVMSYGILLWEIFSLGVNYPYGPVDANFYK 900  
 Db 860 NYVVRGNARLPVKWAPSLFEGYITIKSDVMSYGILLWEIFSLGVNYPYGPVDANFYK 919  
 QY 901 LIQNGFKMDQFYATEEYIIMQSCWAFDSKRPSFPNLTSLFLGQLADABEAMYNQVNDG 960  
 Db 920 LIQNGFKMDQFYATEEYIIMQSCWAFDSKRPSFPNLTSLFLGQLADABEAMYNQVNDG 979  
 QY 961 RVSECPHTYQNRPFPSREMDLGLLSPOAQVEDS 993  
 Db 980 RVSECPHTYQNRPFPSREMDLGLLSPOAQVEDS 1012

RESULT 13  
 AAW63588  
 ID AAW63588 standard; Protein; 983 AA.

XX AAW63588;  
 AC 12-OCT-1998 (first entry)  
 DT Human receptor type protein kinase FLT3 protein SEQ ID NO:19.  
 DE Human; receptor type protein kinase; FLT3; diagnosis; leukaemia;  
 KW membrane-vicinal region; tyrosine kinase; juxtamembrane region.  
 XX Homo sapiens.  
 OS WO9817808-A1.  
 PN 30-APR-1998.  
 PD 13-OCT-1997; 97MO-JP03667.  
 PF 18-OCT-1996; 96JP-0297329.  
 PR (TAKI ) TAKARA SHUZO CO LTD.  
 PA Yokota S;  
 PI WPI; 1998-362333/31.  
 PP N-PSDB; AAV39041.  
 DR Nucleic acid sequences encoding receptor type protein kinase -  
 PT useful for diagnosing adult acute leukaemia of, e.g. FAB sub-classes  
 PS M2  
 Claim 6; Page 45-50; 80pp; Japanese.  
 XX New nucleic acid sequences have been isolated which encode receptor type  
 CC protein kinases (especially a tyrosine kinases) having tandem repeats in  
 CC the juxtamembrane region. Also described in the present invention are:  
 CC (1) any one of 10 amino acid (aa) sequences (AAW63580 to AAW63589)  
 CC representing FLT3 juxtamembrane receptor type protein kinases found in  
 CC leukaemia patients of FAB subclass M2, M4 or M5 (nucleic acid sequences  
 CC encoding the proteins are given in AAV19028 to AAV39042); (2) antibodies  
 CC recognising the kinases or their portions including the tandem repeat  
 CC region; (3) nucleic acid sequences hybridising with the nucleic acids;  
 CC (4) a method for detecting the nucleic acid sequences in human tissue  
 CC samples by: (i) extracting nucleic acid; (ii) amplifying the kinase  
 CC gene, and (iii) determining the size of the gene for comparing the size  
 CC of the normal gene not containing tandem repeats, and (5) kits for  
 CC carrying out the detection. The products and methods may be used for  
 CC diagnosing adult acute leukaemia of FAB subclasses M2, M4 and M5.  
 XX Sequence 983 AA;  
 SQ Query Match 97.8%; Score 5159; DB 19; Length 983;  
 Best Local Similarity 99.0%; Pred. No. 0;  
 Matches 973; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
 QY 1 MPALARDAGTVPLLVFSAMIFGTITNODLPVVKVLIHKNKNDSSVGKSSSYPMVSESP 60  
 Db 1 MPALARDAGTVPLLVFSAMIFGTITNODLPVVKVLIHKNKNDSSVGKSSSYPMVSESP 60  
 QY 61 EDLGCALRPOSSGTVYEAFAVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCQPHFDL 120  
 Db 61 EDLGCALRPOSSGTVYEAFAVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCQPHFDL 120  
 QY 121 QNRGVSVMLKMTETQAGEYLLFIQSEATNTYILFTVSVIRNTLLYTLRRPYFRKMENQD 180  
 Db 121 QNRGVSVMLKMTETQAGEYLLFIQSEATNTYILFTVSVIRNTLLYTLRRPYFRKMENQD 180  
 QY 181 ALVCISSEVPPIVEWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240  
 Db 181 ALVCISSEVPPIVEWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240  
 QY 241 CTRLFTIDLNOTPOTTLFQPLFKVGEPLWIRCAVHVNHGFGLTWELNKALBEGNYFEM 300

Db 241 CTRLFTIDLNOTPQTTLPOLFLKVEPLWIRCKAVHVNHGFLTWELNKALEEGNYFEM 300  
 Qy 301 STYSTNRTMIRILFAFVSSVARNDGYTTCSSSKHPSQSALVTIYKGFINATNSSEDEYE 360  
 Db 301 STYSTNRTMIRILFAFVSSVARNDGYTTCSSSKHPSQSALVTIYKGFINATNSSEDEYE 360  
 Qy 361 IDQYEEFCFVRKAYPQIRCTWTFSRKSFPCEQKGLDNGYSISKFCNKHQPGYIIFHA 420  
 Db 361 IDQYEEFCFVRKAYPQIRCTWTFSRKSFPCEQKGLDNGYSISKFCNKHQPGYIIFHA 420  
 Qy 421 ENDDAQFTKMTLNIRRRKQVLAESAASQSCFSDGYPLPSMTWKCKSDKSPNCTEITE 480  
 Db 421 ENDDAQFTKMTLNIRRRKQVLAESAASQSCFSDGYPLPSMTWKCKSDKSPNCTEITE 480  
 Qy 481 GVMNRKANRKFQGVSSSTLNMSBAIKGFLVKCCAYNSLGTSCETILLNSPGPPFFIQD 540  
 Db 481 GVMNRKANRKFQGVSSSTLNMSBAIKGFLVKCCAYNSLGTSCETILLNSPGPPFFIQD 540  
 Qy 541 NISFYATIGVCLLFIVVLTLLIICHYKQKQFRYESQIQMVQVTGSSDNEYFYVDFREYED 600  
 Db 541 NISFYATIGVCLLFIVVLTLLIICHYKQKQFRYESQIQMVQVTGSSDNEYFYVDFREYED 600  
 Qy 601 L-----KWEPPRENLEFGKVLGSGAFCKVMNATAYGISKTGVSIOQAVKMLKEKA 650  
 Db 601 LKWEPPRENLEFGKVLGSGAFCKVMNATAYGISKTGVSIOQAVKMLKEKA 650  
 Qy 651 DSSREALMSLKMWTQGLSHENIVNLLGACTLSGPIYLIIFYCCYGDLLNLYLRKREKF 710  
 Db 651 DSSREALMSLKMWTQGLSHENIVNLLGACTLSGPIYLIIFYCCYGDLLNLYLRKREKF 710  
 Qy 711 HRTWTEIFKEHNFSYPTFQSHPNSSMPGSRVQIHPDSQISGLHGNFSHSEDEIEYEN 770  
 Db 721 HRTWTEIFKEHNFSYPTFQSHPNSSMPGSRVQIHPDSQISGLHGNFSHSEDEIEYEN 780  
 Qy 771 QKRLBEEDNLVLTEDILLCFAYQVAKGMFELEFKSCVHRDLAARNVLVTHGKVKICDF 830  
 Db 781 QKRLBEEDNLVLTEDILLCFAYQVAKGMFELEFKSCVHRDLAARNVLVTHGKVKICDF 840  
 Qy 831 GLARDIMSDSNVVRGNARLPVKWAPESLFGIYTIKSDVMSYIGILLWEIIFSLGVNYPY 890  
 Db 841 GLARDIMSDSNVVRGNARLPVKWAPESLFGIYTIKSDVMSYIGILLWEIIFSLGVNYPY 900  
 Qy 891 GIPVDANFYKLIQNGFKMDQPYATEEIIYIMQSCWAFDSRKRPPNLTSLFGCOLADA 950  
 Db 901 GIPVDANFYKLIQNGFKMDQPYATEEIIYIMQSCWAFDSRKRPPNLTSLFGCOLADA 960  
 Qy 951 EAMYNQVNDGRVSECPHTYQNER 973  
 Db 961 EAMYNQVNDGRVSECPHTYQNER 983

RESULT 14  
 AAW63587  
 ID AAW63587 standard; Protein; 986 AA.  
 AC  
 XX AAW63587;  
 DT 12-OCT-1998 (first entry)  
 XX  
 DE Human receptor type protein kinase FLT3 protein SEQ ID NO:18.  
 XX  
 KW Human; receptor type protein kinase; FLT3; diagnosis; leukaemia;  
 KW membrane-vicinal region; tyrosine kinase; juxtamembrane region.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9817808-A1.  
 XX  
 PD 30-APR-1998.  
 XX  
 PF 13-OCT-1997; 97WO-JP03667.  
 XX  
 PR 18-OCT-1996; 96JP-0297329.

XX (TAKI ) TAKARA SHUZO CO LTD.  
 PA Yokota S;  
 XX WPI; 1998-362333/31.  
 DR N-PSDB; AAV39040.  
 XX  
 PT Nucleic acid sequences encoding receptor type protein kinase -  
 useful for diagnosing adult acute leukaemia of, e.g. FAB sub-classes  
 M2  
 PT  
 PS Claim 6; Page 40-45; 80pp; Japanese.  
 XX  
 CC New nucleic acid sequences have been isolated which encode receptor type  
 protein kinases (especially a tyrosine kinases) having tandem repeats in  
 the juxtamembrane region. Also described in the present invention are:  
 CC (1) any one of 10 amino acid (aa) sequences (AAW63580 to AAW63589)  
 CC representing FLT3 juxtamembrane receptor type protein kinases found in  
 CC leukaemia patients of FAB subclass M2, M4 or M5 (nucleic acid sequences  
 CC encoding the proteins are given in AAV39028 to AAV39042); (2) antibodies  
 CC recognising the kinases or their portions including the tandem repeat  
 CC region; (3) nucleic acid sequences hybridising with the nucleic acids;  
 CC (4) a method for detecting the nucleic acid sequences in human tissue  
 CC samples by: (i) extracting nucleic acid; (ii) amplifying the kinase  
 CC gene, and (iii) determining the size of the gene for comparing the size  
 CC of the normal gene not containing tandem repeats, and (5) kits for  
 CC carrying out the detection. The products and methods may be used for  
 CC diagnosing adult acute leukaemia of FAB subclasses M2, M4 and M5.  
 XX Sequence 986 AA;  
 SQ  
 Query Match 97.8%; Score 5157.5; DB 19; Length 986;  
 Best Local Similarity 98.7%; Pred. No. 0;  
 Matches 973; Conservative 0; Mismatches 0; Indels 13; Gaps 1;  
 Qy 1 MPALARDAGTVPLLVFSAMIFGTTNQDLPIVKVLINHKNDSSVGKSSYPMVSESP 60  
 Db 1 MPALARDAGTVPLLVFSAMIFGTTNQDLPIVKVLINHKNDSSVGKSSYPMVSESP 60  
 Qy 61 EDLGCALPQSSGTVEAAAEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCPHFDL 120  
 Db 61 EDLGCALPQSSGTVEAAAEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCPHFDL 120  
 Qy 121 QNRGVSVMLKMTETQAGEYLLFIQSEATNTILFTVSIRNTLLYTLRRPYFRKMNQD 180  
 Db 121 QNRGVSVMLKMTETQAGEYLLFIQSEATNTILFTVSIRNTLLYTLRRPYFRKMNQD 180  
 Qy 181 ALVCISESVPEPIVEWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240  
 Db 181 ALVCISESVPEPIVEWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240  
 Qy 241 CTRLFTIDLNOTPQTTLPOLFLKVEPLWIRCKAVHVNHGFLTWELNKALEEGNYFEM 300  
 Db 241 CTRLFTIDLNOTPQTTLPOLFLKVEPLWIRCKAVHVNHGFLTWELNKALEEGNYFEM 300  
 Qy 301 STYSTNRTMIRILFAFVSSVARNDGYTTCSSSKHPSQSALVTIYKGFINATNSSEDEYE 360  
 Db 301 STYSTNRTMIRILFAFVSSVARNDGYTTCSSSKHPSQSALVTIYKGFINATNSSEDEYE 360  
 Qy 361 IDQYEEFCFVRKAYPQIRCTWTFSRKSFPCEQKGLDNGYSISKFCNKHQPGYIIFHA 420  
 Db 361 IDQYEEFCFVRKAYPQIRCTWTFSRKSFPCEQKGLDNGYSISKFCNKHQPGYIIFHA 420  
 Qy 421 ENDDAQFTKMTLNIRRRKQVLAESAASQSCFSDGYPLPSMTWKCKSDKSPNCTEITE 480  
 Db 421 ENDDAQFTKMTLNIRRRKQVLAESAASQSCFSDGYPLPSMTWKCKSDKSPNCTEITE 480  
 Qy 481 GVMNRKANRKFQGVSSSTLNMSBAIKGFLVKCCAYNSLGTSCETILLNSPGPPFFIQD 540  
 Db 481 GVMNRKANRKFQGVSSSTLNMSBAIKGFLVKCCAYNSLGTSCETILLNSPGPPFFIQD 540  
 Qy 541 NISFYATIGVCLLFIVVLTLLIICHYKQKQFRYESQIQMVQVTGSSDNEYFYVDFREYED 592

Db 541 NISFYATIGVCLLFIVLTLCHYKQFRYESQLQMVQVTGSSDNEYFYVDREYED 600  
 QY 593 -----DPREYEDLKWEFFRENLEFGKVLGSGAGKVMNATAYGISTGYSIQVAVKMLK 647  
 Db 601 LKWEFDPREYEDLKWEFFRENLEFGKVLGSGAGKVMNATAYGISTGYSIQVAVKMLK 660  
 QY 648 EKADSSERREALMSELKMTQLGSHENIVNLLGACTLSGPIYLPEYCCYGDLLNLYLSKR 707  
 Db 661 EKADSSERREALMSELKMTQLGSHENIVNLLGACTLSGPIYLPEYCCYGDLLNLYLSKR 720  
 QY 708 EKPHRTWTETFEKHNFSFYPTFQSHPNSSMPGSRREVQIHPDSQISGLHNSFHSDEIE 767  
 Db 721 EKPHRTWTETFEKHNFSFYPTFQSHPNSSMPGSRREVQIHPDSQISGLHNSFHSDEIE 780  
 QY 768 YENOKLEEEEDLNLTTFEDLLCFAYQVAKGMFELEFKSCVHRDLAARNVLVTHGKVVKI 827  
 Db 781 YENOKLEEEEDLNLTTFEDLLCFAYQVAKGMFELEFKSCVHRDLAARNVLVTHGKVVKI 840  
 QY 828 CDFGLARDIMSDSNVVRGNARLPVKWMAPESELEFEGITYTIKSDVWSYGILLWEIFSLGVN 887  
 Db 841 CDFGLARDIMSDSNVVRGNARLPVKWMAPESELEFEGITYTIKSDVWSYGILLWEIFSLGVN 900  
 QY 888 PYPGIPVDANFYKLIQNGFKMDQFFVATEBIYIIMQSCWAFDSKRKPSFNLTSFLGCQL 947  
 Db 901 PYPGIPVDANFYKLIQNGFKMDQFFVATEBIYIIMQSCWAFDSKRKPSFNLTSFLGCQL 960  
 QY 948 ADAEEAMYQNDGRVSECPHTYQNR 973  
 Db 961 ADAEEAMYQNDGRVSECPHTYQNR 986

RESULT 15

AAW63589  
 ID AAW63589 standard; Protein; 986 AA.  
 AC AAW63589;  
 XX  
 DT 12-OCT-1998 (first entry)  
 XX  
 DE Human receptor type protein kinase FLT3 protein SEQ ID NO:20.  
 KW Human; receptor type protein kinase; FLT3; diagnosis; leukaemia;  
 KW membrane-vicinal region; tyrosine kinase; juxtamembrane region.  
 XX  
 OS Homo sapiens.  
 PN WO9817808-A1.  
 XX  
 PD 30-APR-1998.  
 XX  
 PF 13-OCT-1997; 97WO-03667.  
 XX  
 PR 18-OCT-1996; 96JP-0297329.  
 XX  
 PA (TAKI) TAKARA SHUZO CO LTD.  
 XX  
 PI Yokota S;  
 DR WPI; 1998-362333/31.  
 DR N-PSDB; AAV39042.  
 XX  
 PT Nucleic acid sequences encoding receptor type protein kinase -  
 PT useful for diagnosing adult acute leukaemia of, e.g. FAB sub-classes  
 PT M2  
 XX  
 PS Claim 6; Page 50-55; 80pp; Japanese.  
 XX  
 CC New nucleic acid sequences have been isolated which encode receptor type  
 CC protein kinases (especially a tyrosine kinases) having tandem repeats in  
 CC the juxtamembrane region. Also described in the present invention are:  
 CC (1) any one of 10 amino acid (aa) sequences (AAW63580 to AAW63589)  
 CC representing FLT3 juxtamembrane receptor type protein kinases found in

CC leukaemia patients of FAB subclasses M2, M4 or M5 (nucleic acid sequences  
 CC encoding the proteins are given in AAV39028 to AAV39042); (2) antibodies  
 CC recognising the kinases or their portions including the tandem repeat  
 CC region; (3) nucleic acid sequences hybridising with the nucleic acids;  
 CC (4) a method for detecting the nucleic acid sequences in human tissue  
 CC samples by: (i) extracting nucleic acid; (ii) amplifying the kinase  
 CC gene, and (iii) determining the size of the gene for comparing the size  
 CC of the normal gene not containing tandem repeats, and (5) kits for  
 CC carrying out the detection. The products and methods may be used for  
 CC diagnosing adult acute leukaemia of FAB subclasses M2, M4 and M5.  
 XX

SQ Sequence 986 AA;

Query Match 97.8%; Score 5157.5; DB 19; Length 986;  
 Best Local Similarity 98.74; Pred. No. 0;  
 Matches 973; Conservative 0; Mismatches 0; Indels 13; Gaps 1;

QY 1 MPALARDAGTVPVLLVFSAMIFGTITNODLPVTKCVLINHKNDSSVSKSSYPMVSESP 60  
 Db 1 MPALARDAGTVPVLLVFSAMIFGTITNODLPVTKCVLINHKNDSSVSKSSYPMVSESP 60  
 QY 61 EDLGCALRPOSSGTVEAAAEVDVSASITLQVLVDAPGNISCLWVFKHSLNCPHFDL 120  
 Db 61 EDLGCALRPOSSGTVEAAAEVDVSASITLQVLVDAPGNISCLWVFKHSLNCPHFDL 120  
 QY 121 QNRGVSMVLKMTETQAGEYLLFIQSEATNYTLFTVSIINTLLYTLRRPYFRKMNQD 180  
 Db 121 QNRGVSMVLKMTETQAGEYLLFIQSEATNYTLFTVSIINTLLYTLRRPYFRKMNQD 180  
 QY 181 ALVCISESPEPIVEWVLCDSQESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240  
 Db 181 ALVCISESPEPIVEWVLCDSQESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240  
 QY 241 CTRLFTIDLNOTPTTLPOLFLKVGEPMLTRCAVHVNHGFGLTWELENKALEEGNFEM 300  
 Db 241 CTRLFTIDLNOTPTTLPOLFLKVGEPMLTRCAVHVNHGFGLTWELENKALEEGNFEM 300  
 QY 301 STYSTNRTMIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFNATNSSDYE 360  
 Db 301 STYSTNRTMIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFNATNSSDYE 360  
 QY 361 IDQEEFCFSVRPKAYPOIRCTWTFSRKSPPCQKGLDNGYSISKFCNKHQPGEYIFHA 420  
 Db 361 IDQEEFCFSVRPKAYPOIRCTWTFSRKSPPCQKGLDNGYSISKFCNKHQPGEYIFHA 420  
 QY 421 ENDDAQTMTFTLNRKQVLAESAASQASCFSDGYPLPSWTWKKCDSPNCTEITE 480  
 Db 421 ENDDAQTMTFTLNRKQVLAESAASQASCFSDGYPLPSWTWKKCDSPNCTEITE 480  
 QY 481 GVMNRKANERKVFQWVSSSTLNWSEAIKGLVKCCAYNSLGTSCETILLNSPGPFPIQD 540  
 Db 481 GVMNRKANERKVFQWVSSSTLNWSEAIKGLVKCCAYNSLGTSCETILLNSPGPFPIQD 540  
 QY 541 NISFYATIGVCLLFIVLTLCHYKQFRYESQLQMVQVTGSSDNEYFYVDREYED 587  
 Db 541 NISFYATIGVCLLFIVLTLCHYKQFRYESQLQMVQVTGSSDNEYFYVDREYED 600  
 QY 588 EYFYVDPREYEDLKWEFFRENLEFGKVLGSGAGKVMNATAYGISTGYSIQVAVKMLK 647  
 Db 601 EYFYVDPREYEDLKWEFFRENLEFGKVLGSGAGKVMNATAYGISTGYSIQVAVKMLK 660  
 QY 648 EKADSSERREALMSELKMTQLGSHENIVNLLGACTLSGPIYLFIYCCYGDLLNLYLSKR 707  
 Db 661 EKADSSERREALMSELKMTQLGSHENIVNLLGACTLSGPIYLFIYCCYGDLLNLYLSKR 720  
 QY 708 EKPHRTWTETFEKHNFSFYPTFQSHPNSSMPGSRREVQIHPDSQISGLHNSFHSDEIE 767  
 Db 721 EKPHRTWTETFEKHNFSFYPTFQSHPNSSMPGSRREVQIHPDSQISGLHNSFHSDEIE 780  
 QY 768 YENOKLEEEEDLNLTTFEDLLCFAYQVAKGMFELEFKSCVHRDLAARNVLVTHGKVVKI 827  
 Db 781 YENOKLEEEEDLNLTTFEDLLCFAYQVAKGMFELEFKSCVHRDLAARNVLVTHGKVVKI 840

QY	828	CDFGIARDIMSDSNVVRGNARLPVKWMAPELSLFEIGIYTIKSDVWSYGILLWEIFSLGVN	887
Db	841	CDFGIARDIMSDSNVVRGNARLPVKWMAPELSLFEIGIYTIKSDVWSYGILLWEIFSLGVN	900
QY	888	PYPGIPVDANFYKLIQNGFKMDOPFYATEEIIYIMQSCWAFDSRKRPSPNLTSLGCOL	947
Db	901	PYPGIPVDANFYKLIQNGFKMDOPFYATEEIIYIMQSCWAFDSRKRPSPNLTSLGCOL	960
QY	948	ADAEEMQYONVDGRVSECPHTYQNR	973
Db	961	ADAEEMQYONVDGRVSECPHTYQNR	986

Search completed: August 26, 2003, 07:24:13  
Job time : 66.5315 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2003, 07:27:59 ; Search time 45.5229 Seconds  
(without alignments)  
2878.506 Million cell updates/sec

Title: US-09-919-408A-4  
Perfect score: 5274  
Sequence: 1 MPALARDAGTVLLVVFSA.....PFSREMDLGLLSPPQAQVEDS 993

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 497079 seqs, 131961718 residues

Total number of hits satisfying chosen parameters: 497079

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : Published Applications AA.\*
- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
  - 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
  - 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
  - 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
  - 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
  - 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
  - 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
  - 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
  - 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
  - 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
  - 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5274	100.0	993	9 US-09-919-408-4	Sequence 4, Appli
2	5274	100.0	993	10 US-09-872-136-4	Sequence 2, Appli
3	4429.5	84.0	992	9 US-09-919-408-2	Sequence 2, Appli
4	4429.5	84.0	992	10 US-09-872-136-2	Sequence 2, Appli
5	1266	24.0	975	15 US-10-192-867-2	Sequence 32, Appli
6	1251	23.7	976	14 US-10-099-895-32	Sequence 4, Appli
7	1251	23.7	976	15 US-10-192-867-4	Sequence 10, Appli
8	1216.5	23.1	972	10 US-09-944-807-10	Sequence 4, Appli
9	1166.5	22.1	1088	11 US-09-961-403-4	Sequence 2, Appli
10	1166	22.1	1089	9 US-09-769-987-2	Sequence 90, Appli
11	1166	22.1	1089	10 US-09-919-497-90	Sequence 2, Appli
12	1166	22.1	1089	10 US-09-866-510-2	Sequence 36, Appli
13	1166	22.1	1089	10 US-09-955-363-36	Sequence 10, Appli
14	1163	22.1	1089	10 US-09-866-510-10	Sequence 4, Appli
15	1162	22.0	1089	10 US-09-866-510-4	

16	1161	22.0	1089	10 US-09-866-510-8	Sequence 8, Appli
17	1160	22.0	1089	10 US-09-866-510-6	Sequence 6, Appli
18	1086.5	20.6	1090	10 US-09-866-510-14	Sequence 14, Appli
19	1086.5	20.6	1106	10 US-09-866-510-22	Sequence 22, Appli
20	1086.5	20.6	1106	10 US-09-955-363-2	Sequence 2, Appli
21	1082.5	20.5	1106	10 US-09-866-510-16	Sequence 16, Appli
22	1081.5	20.5	1106	10 US-09-866-510-20	Sequence 20, Appli
23	1080.5	20.5	1106	10 US-09-866-510-18	Sequence 18, Appli
24	990	18.8	1338	12 US-10-021-660-84	Sequence 84, Appli
25	990	18.8	1338	15 US-10-039-585-44	Sequence 44, Appli
26	957	18.1	386	9 US-09-939-754-6	Sequence 6, Appli
27	957	18.1	386	10 US-09-939-832-6	Sequence 6, Appli
28	957	18.1	386	10 US-09-939-833-6	Sequence 2, Appli
29	951.5	18.0	1356	15 US-10-022-939-2	Sequence 2, Appli
30	951.5	18.0	1356	15 US-10-100-405A-2	Sequence 2, Appli
31	950.5	18.0	1356	11 US-09-969-537-7	Sequence 7, Appli
32	944.5	17.9	1354	15 US-10-262-538-30	Sequence 30, Appli
33	934	17.7	1367	9 US-09-766-678-2	Sequence 2, Appli
34	923	17.5	1367	9 US-09-919-408-6	Sequence 6, Appli
35	923	17.5	1367	10 US-09-872-136-6	Sequence 6, Appli
36	907.5	17.2	1363	11 US-09-375-248-19	Sequence 19, Appli
37	904.5	17.2	1298	10 US-09-982-610-33	Sequence 33, Appli
38	904.5	17.2	1298	15 US-10-262-538-32	Sequence 32, Appli
39	904.5	17.2	1363	11 US-09-375-248-2	Sequence 2, Appli
40	902.5	17.1	1368	15 US-10-105-901-34	Sequence 34, Appli
41	902	17.1	1362	15 US-10-105-901-33	Sequence 33, Appli
42	880.5	16.7	1363	15 US-10-105-901-32	Sequence 32, Appli
43	859.5	16.3	367	9 US-09-939-754-9	Sequence 9, Appli
44	859.5	16.3	367	10 US-09-939-832-9	Sequence 9, Appli
45	859.5	16.3	367	10 US-09-939-833-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1  
US-09-919-408-4  
; Sequence 4, Application US/09919408  
; Patent No. US20020072077A1  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ImClone Systems Incorporated  
; STREET: 180 Varick Street  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/09/919,408  
; APPLICATION NUMBER: US/09/919,408  
; FILING DATE: 31-Jul-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; FILING DATE: 07/977,451  
; APPLICATION NUMBER: <Unknown>  
; FILING DATE: 07/906,397  
; APPLICATION NUMBER: US PCT/US92/05401  
; FILING DATE: 26-JUN-1992  
; APPLICATION NUMBER: US PCT/US92/02750  
; FILING DATE: 15-APR-1992  
; APPLICATION NUMBER: US PCT/US92/02750  
; FILING DATE: 02-APR-1992  
; APPLICATION NUMBER: US 07/813,593  
; FILING DATE: 24-DEC-1991



APPLICATION NUMBER: US 07/793,065  
FILING DATE: 15-NOV-1991  
APPLICATION NUMBER: US 07/728,913  
FILING DATE: 28-JUN-1991  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Felt, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 993 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-919-408-4

Query Match 100.0%; Score 5274; DB 9; Length 993;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPLLVVFSAIFGTITNQLDLPVKVLINHKNDSSVGKSSYPVMVSESP 60  
DB 1 MPALARDAGTVPLLVVFSAIFGTITNQLDLPVKVLINHKNDSSVGKSSYPVMVSESP 60  
QY 61 EDLGCALRPQSSGTGYEAAAEVDVSASITLQVLVDAPGNISCLMVLVFKHSSINCOPHDL 120  
DB 61 EDLGCALRPQSSGTGYEAAAEVDVSASITLQVLVDAPGNISCLMVLVFKHSSINCOPHDL 120  
QY 121 QNRGVVSVILKWTQAGEYLLFTQSEATNTYLLFTVSIKNTLYTLRPPYFRKMNOD 180  
DB 121 QNRGVVSVILKWTQAGEYLLFTQSEATNTYLLFTVSIKNTLYTLRPPYFRKMNOD 180  
QY 181 ALVCISESVPEIVNVLCDQGESCKEESPAVVKKEKVLHFLGTDIRCCARNELGRE 240  
DB 181 ALVCISESVPEIVNVLCDQGESCKEESPAVVKKEKVLHFLGTDIRCCARNELGRE 240  
QY 241 CTRLTIDLNQPTTLPOLFLKVEPLMIRKAVHNVHFGPLTWELNKALBEGNYFEM 300  
DB 241 CTRLTIDLNQPTTLPOLFLKVEPLMIRKAVHNVHFGPLTWELNKALBEGNYFEM 300  
QY 301 STYSTNRTMIRILFAVSSVARNDTGYTCSKSKHPSQSALVTIVGKFINATNSSEYD 360  
DB 301 STYSTNRTMIRILFAVSSVARNDTGYTCSKSKHPSQSALVTIVGKFINATNSSEYD 360  
QY 361 IDQYEEFCFSVRFKAYPQIRCTWTFSRKSFPCEQKGLDNGYSISKFCNKHQPGEYIFHA 420  
DB 361 IDQYEEFCFSVRFKAYPQIRCTWTFSRKSFPCEQKGLDNGYSISKFCNKHQPGEYIFHA 420  
QY 421 ENDDAQFTMTINIRKQVLAESAASQAFSDGYPLPSWTWKCKSDKSPNCTEITE 480  
DB 421 ENDDAQFTMTINIRKQVLAESAASQAFSDGYPLPSWTWKCKSDKSPNCTEITE 480  
QY 481 GYVNRKANRKFQGVVSSSTLNNSAIGFLVKCCAYNSLGTSCETILLNSGPPFFIOD 540  
DB 481 GYVNRKANRKFQGVVSSSTLNNSAIGFLVKCCAYNSLGTSCETILLNSGPPFFIOD 540  
QY 541 NISFVATIGVCLLFIIVLTLCHKYKQFRVYESQLQVQVGTGSSDNEYFYVDFREYED 600  
DB 541 NISFVATIGVCLLFIIVLTLCHKYKQFRVYESQLQVQVGTGSSDNEYFYVDFREYED 600  
QY 601 LKWEPPRENLEFGKVLGSAFGKVMNATAYGISTGVSIQVAVKMLKEKADSSEREA 660  
DB 601 LKWEPPRENLEFGKVLGSAFGKVMNATAYGISTGVSIQVAVKMLKEKADSSEREA 660  
QY 661 ELKQMTQLGSHENIVNLLGACTLSGPIYLIFFCYCCYGDLLNLYRSREKPHRTWTF 720  
DB 661 ELKQMTQLGSHENIVNLLGACTLSGPIYLIFFCYCCYGDLLNLYRSREKPHRTWTF 720

QY 721 HNFSPYPTFQSHPNSSMPCSRREVQIHPDSQDQISGLHGSHFSEDEIEYENQKRELEEDL 780  
DB 721 HNFSPYPTFQSHPNSSMPCSRREVQIHPDSQDQISGLHGSHFSEDEIEYENQKRELEEDL 780  
QY 781 NVLTPEDLCPAYQVAKGMEPLFKPCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDS 840  
DB 781 NVLTPEDLCPAYQVAKGMEPLFKPCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDS 840  
QY 841 NYVVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIFSLGVPNPYPGIPVDANFYK 900  
DB 841 NYVVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIFSLGVPNPYPGIPVDANFYK 900  
QY 901 LIQGFKNQDOPFYATEEYIIMQSCWAFDSRKRPSFPNLTSLGCOLADAEAEAMYQNVDG 960  
DB 901 LIQGFKNQDOPFYATEEYIIMQSCWAFDSRKRPSFPNLTSLGCOLADAEAEAMYQNVDG 960  
QY 961 RVSECPHYQNRPPFSREMDLGLLSPQAQVEDS 993  
DB 961 RVSECPHYQNRPPFSREMDLGLLSPQAQVEDS 993

RESULT 2  
US-09-872-136-4  
; Sequence 4, Application US/09872136  
; Patent No. US20020119545A1  
; GENERAL INFORMATION:  
; APPLICANT: Lewischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; ; RECEPTORS AND THEIR LIGANDS  
; ;  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ImClone Systems Incorporated  
; STREET: 180 Varick Street  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/872,136  
; FILING DATE: 01-Jun-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/208,786  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: US/09/021,324  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: US/07/977,451  
; FILING DATE: 1992-11-19  
; APPLICATION NUMBER: US 07/906,397  
; FILING DATE: 26-JUN-1992  
; APPLICATION NUMBER: US PCT/US92/05401  
; FILING DATE: 26-JUN-1992  
; APPLICATION NUMBER: TW 81102961  
; FILING DATE: 15-APR-1992  
; APPLICATION NUMBER: US PCT/US92/02750  
; FILING DATE: 02-APR-1992  
; APPLICATION NUMBER: US 07/813,593  
; FILING DATE: 24-DEC-1991  
; APPLICATION NUMBER: US 07/793,065  
; FILING DATE: 15-NOV-1991  
; APPLICATION NUMBER: US 07/728,913  
; FILING DATE: 28-JUN-1991  
; APPLICATION NUMBER: US 07/679,666  
; FILING DATE: 02-APR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Felt, Irving N.  
; REGISTRATION NUMBER: 28,601

REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 993 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-872-136-4

Query Match 100.0%; Score 5274; DB 10; Length 993;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPVLLVFSAMIFGTITNODLPVVKVLIHKNHNDSSVKGSSSYPMVSESP 60  
Db 1 MPALARDAGTVPVLLVFSAMIFGTITNODLPVVKVLIHKNHNDSSVKGSSSYPMVSESP 60

QY 61 EDLGCALRPQSSGTVEAAAEVDVVSASITLQVLVDAPGNISCLWVPKHSLSLNCQPHFDL 120  
Db 61 EDLGCALRPQSSGTVEAAAEVDVVSASITLQVLVDAPGNISCLWVPKHSLSLNCQPHFDL 120

QY 121 QNRGVVSWILKMTETQAGEYLLFIOSEATNTILFTVSRNTLLYTLRPPYKRMENQD 180  
Db 121 QNRGVVSWILKMTETQAGEYLLFIOSEATNTILFTVSRNTLLYTLRPPYKRMENQD 180

QY 181 ALVCISESPEPIVEVWVLCDSQGESCKEESPAVVKEEKVHLHFGTDIRCCARNELGRE 240  
Db 181 ALVCISESPEPIVEVWVLCDSQGESCKEESPAVVKEEKVHLHFGTDIRCCARNELGRE 240

QY 241 CTRLFTIDLNTQPTTLPOLFLKVGEPWLIRCAVHNVHGFGLTWELNKALBEGNYFEM 300  
Db 241 CTRLFTIDLNTQPTTLPOLFLKVGEPWLIRCAVHNVHGFGLTWELNKALBEGNYFEM 300

QY 301 STYSTNRTMIRILFAVSSVARNDTGYYTSSSKHPSQSALVTIVGKGFNATNSSEDEYE 360  
Db 301 STYSTNRTMIRILFAVSSVARNDTGYYTSSSKHPSQSALVTIVGKGFNATNSSEDEYE 360

QY 361 IDQEEFCFSVRPKAYPQIRCTWTFGRKSPPCBQKGLDNGYSISKCNHKGQGEYIFHA 420  
Db 361 IDQEEFCFSVRPKAYPQIRCTWTFGRKSPPCBQKGLDNGYSISKCNHKGQGEYIFHA 420

QY 421 ENDDAQFTKFTLNIRKQVLAESAASQSCFSDGYPPLSWTWKCKSDKSPNCTEITE 480  
Db 421 ENDDAQFTKFTLNIRKQVLAESAASQSCFSDGYPPLSWTWKCKSDKSPNCTEITE 480

QY 481 GWNRRKANRKFQGWVSSSTLANSEAIKGPLVKCCAYNSLGTSCETILLNSPGPPPIQD 540  
Db 481 GWNRRKANRKFQGWVSSSTLANSEAIKGPLVKCCAYNSLGTSCETILLNSPGPPPIQD 540

QY 541 NISFYATIGVCLLFIIVLLILLICHYKQPRYESQLQVQVGTSSDNEFYVDFREYVD 600  
Db 541 NISFYATIGVCLLFIIVLLILLICHYKQPRYESQLQVQVGTSSDNEFYVDFREYVD 600

QY 601 LKWEPPRENLEFGKVLGSGAFGKVMNATAYIGISKTGVSIGQAVKMLKEKADSSEREAALS 660  
Db 601 LKWEPPRENLEFGKVLGSGAFGKVMNATAYIGISKTGVSIGQAVKMLKEKADSSEREAALS 660

QY 661 ELKMMTQLGSHENIVNLLGACTLSGPIYLIPEYCCYGDLLNLYRSRKRFHRTWTPIKPE 720  
Db 661 ELKMMTQLGSHENIVNLLGACTLSGPIYLIPEYCCYGDLLNLYRSRKRFHRTWTPIKPE 720

QY 721 HNFSPYPTQSPHNSMPCSRVQIHPDSQDLSGLHNSFHSDETEYENQKELEEDL 780  
Db 721 HNFSPYPTQSPHNSMPCSRVQIHPDSQDLSGLHNSFHSDETEYENQKELEEDL 780

QY 781 NVLTTFEDLLCFAYQVAKGNFLEFKSCVHRDLAARNVLVTHGKVKVLCDFGLARDIMSDS 840  
Db 781 NVLTTFEDLLCFAYQVAKGNFLEFKSCVHRDLAARNVLVTHGKVKVLCDFGLARDIMSDS 840

QY 841 NYVVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIFSLGVNYPGIPVDANFYK 900  
Db 841 NYVVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIFSLGVNYPGIPVDANFYK 900

QY 901 LIQNGFKMDQPFYATEEIIYIMQSCWAFDRKRPSFNLTSFLGCOLADAEAMYNQVNDG 960  
Db 901 LIQNGFKMDQPFYATEEIIYIMQSCWAFDRKRPSFNLTSFLGCOLADAEAMYNQVNDG 960

QY 961 RVSECPHTYQNRPRPFSREMDLGLLSPOAQVEDS 993  
Db 961 RVSECPHTYQNRPRPFSREMDLGLLSPOAQVEDS 993

RESULT 3  
US-09-919-408-2  
; Sequence 2, Application US/09919408  
; Patent No. US20020072077A1  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ImClone Systems Incorporated  
; STREET: 180 Varick Street  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/919,408  
; FILING DATE: 31-Jul-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/977,451  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: US 07/906,397  
; FILING DATE: 26-JUN-1992  
; APPLICATION NUMBER: US PCT/US92/05401  
; FILING DATE: 26-JUN-1992  
; APPLICATION NUMBER: TW 81102961  
; FILING DATE: 15-APR-1992  
; APPLICATION NUMBER: US PCT/US92/02750  
; FILING DATE: 02-APR-1992  
; APPLICATION NUMBER: US 07/813,593  
; FILING DATE: 24-DEC-1991  
; APPLICATION NUMBER: US 07/793,065  
; FILING DATE: 15-NOV-1991  
; APPLICATION NUMBER: US 07/728,913  
; FILING DATE: 28-JUN-1991  
; APPLICATION NUMBER: US 07/679,666  
; FILING DATE: 02-APR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Feit, Irving N.  
; REGISTRATION NUMBER: 28,601  
; REFERENCE/DOCKET NUMBER: LEM-3-7P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-645-1405  
; TELEFAX: 212-645-2054  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 992 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-919-408-2

Query Match		84.0%;	Score 4429.5;	DB 9;	Length 992;
Best Local Similarity		84.1%;	Pred. No. 0;		
Matches 836;		Conservative 57;	Mismatches 90;	Indels 11;	Gaps 4;
Qy	1	MPALA-RDAGTVPLLVVFSAMIFGTTINQDLPIKVLINHKNDSSVGKSSSYPMVSES	59		
Db	1	MPALAQRSDRLLLLVLSWILETVNQDLPIKVLISHENNSSAGKPSYRMVRS	60		
Qy	60	PEDLGCALRPQSGTVYEAFAVEVDVSASITLQVLVDAPGNISCLVFKHSSLNCPHF	119		
Db	61	PEDLQCTPRQSGTVYEAFAVEAESSITLQVLATPDLSCLVFKHSSLGCPHF	120		
Qy	120	LQNRGVSVVILKMTQAGEYLLFTQSEATNTILFTYSIRNTILYTLRRPFRKMNQ	179		
Db	121	LQNRGIVSMAILNVTQAGEYLLHTQSERANYTLFTVNRDTQLYVLRPFRKMNQ	180		
Qy	180	DALVCISESPPIVEMVLCDSQGESKESPAVKKEKVLHELFGTDIRCCARNELGR	239		
Db	181	DALLCISEGVPETVEWVLCSSHRESCKECPAVRKEKVLHELFGTDIRCCARNALGR	240		
Qy	240	ECTRLFTIDLNOTPTTLPLQLFKVGEPLWIRCKAVHVNHGFLTWELNKALEEGNYE	299		
Db	241	ECTKLTIDLNAQPSQLPLQLFKVGEPLWIRCKAIHVNHGFLTWELDKALEEGSYE	300		
Qy	300	MSYSTNRTMIRILFAPVSVARNDGYTCSSKHPQSALVTIVCKGPHNATNSSEY	359		
Db	301	MSYSTNRTMIRILLAFVSSVGRNDGYTCSSKHPQSALVTILEKGFINATSSQEEY	360		
Qy	360	EIDQIEECFVSFRKAYPOIRCTWTESRKSFPCEOKGLDNGYSISKFNHKGHPGEYIP	419		
Db	361	EIDPYEKFVSFRKAYPRIRCTWTISQASFPCEQRLGLEDGYISRFPCHNKPGEYIFY	420		
Qy	420	AENDDAQFTKMTFLNIRRPQVLAESAASQSCFSDGYPLPSTWKKCDKSPNCTEEIT	479		
Db	421	AENDDAQFTKMTFLNIRKRPQVLANASASQSCSDGYPLPSTWKKCDKSPNCTEIP	480		
Qy	480	EGVNRKANRKVFQWVSSSTLNSBAIKFLVKCAYNLSLGTSCETILLNSGPPPIQ	539		
Db	481	EGVNRKANRKVFQWVSSSTLNSBAGKLLVKCAYNLSMGTSCTIFLNSGPPPIQ	540		
Qy	540	DNISFYATIGVCLLFTIVLTLCHKYKQFRYESQLOMVOVGTSSDNEYFYVDREY	599		
Db	541	DNISFYATIGLCLPFIIVLVILCHKYKQFRYESQLOMVOVGTSSDNEYFYVDREY	600		
Qy	600	DLKWEPRENLEKGVLSGAFGKVMNATAYGSKTGVSIQVAVKMLKXADSSREALM	659		
Db	601	DLKWEPRENLEKGVLSGAFGVMNATAYGSKTGVSIQVAVKMLKXADSCKEALM	660		
Qy	660	SELKMTQLGSHENIVNLLGACTLSGPIYLIPEYCCYGDLLNVLRSKREKPHRTWTEIPK	719		
Db	661	SELKMTHLGHDNIVNLLGACTLSGPIYLIPEYCCYGDLLNVLRSKREKPHRTWTEIPK	720		
Qy	720	EHNFSYPTFQSHNSMPSGREVQHPDSDQISGLHNSFHSDELEYENOKEL--EEE	777		
Db	721	EHNFSYPTFQAHNSMPSGREVQLHPDQISGLHNSFHSDELEYENOKELAE	780		
Qy	778	EDNLVITFEDLLCFAYOVAKMEFLEPKSCVHRDLAARNVLVTHGKVVKICDFGLARDIM	837		
Db	781	EDNLVITFEDLLCFAYOVAKMEFLEPKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL	840		
Qy	838	SDSNVYVRGNARLPVKWMAPELIFEGYITIKSDVMSYGILLWEIFSLGVNYPGIPVDAN	897		
Db	841	SDSNVYVRGNARLPVKWMAPELIFEGYITIKSDVMSYGILLWEIFSLGVNYPGIPVDAN	900		
Qy	898	FYKLIQNGFKMDQPPYATEIYIIMQSCWAFDSKRPFLNLSPLGLCOLADAEAMQN	957		
Db	901	FYKLIQSGFKMEQFPYATEIYIFVMQSCWAFDSKRPFLNLSPLGLCOLAEAEAC--	957		
Qy	958	VDGRVSECPHYQNRPRFSEMDLGLSPQAQVE	991		
Db	958	-----IRTSIHLPKQAAPQQRG-GLRAQSPQRQVK	986		

RESULT 4  
US-09-872-136-2  
; Sequence 2, Application US/09872136  
; Patent No. US20020119545A1  
GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ibor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ImClone Systems Incorporated  
; STREET: 180 Varick Street  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10014  
COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/872.136  
; FILING DATE: 01-Jun-2001  
; CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/208,786  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: US/09/021,324  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: US/07/977,451  
; FILING DATE: 1992-11-19  
; APPLICATION NUMBER: US 07/906,397  
; FILING DATE: 26-JUN-1992  
; APPLICATION NUMBER: US PCT/US92/05401  
; FILING DATE: 26-JUN-1992  
; APPLICATION NUMBER: TW 81102961  
; FILING DATE: 15-APR-1992  
; APPLICATION NUMBER: US PCT/US92/02750  
; FILING DATE: 02-APR-1992  
; APPLICATION NUMBER: US 07/813,593  
; FILING DATE: 24-DEC-1991  
; APPLICATION NUMBER: US 07/793,065  
; FILING DATE: 15-NOV-1991  
; APPLICATION NUMBER: US 07/728,913  
; FILING DATE: 28-JUN-1991  
; APPLICATION NUMBER: US 07/679,666  
; FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
; NAME: Feit, Irving N.  
; REGISTRATION NUMBER: 28,601  
; REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-645-1405  
; TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 992 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-872-136-2

Query Match		84.0%;	Score 4429.5;	DB 10;	Length 992;
Best Local Similarity		84.1%;	Pred. No. 0;		
Matches 836;		Conservative 57;	Mismatches 90;	Indels 11;	Gaps 4;
Qy	1	MPALA-RDAGTVPLLVVFSAMIFGTTINQDLPIKVLINHKNDSSVGKSSSYPMVSES	59		
Db	1	MPALAQRSDRLLLLVLSWILETVNQDLPIKVLISHENNSSAGKPSYRMVRS	60		
Qy	60	PEDLGCALRPQSGTVYEAFAVEVDVSASITLQVLVDAPGNISCLVFKHSSLNCPHF	119		

```

Db 61 PEDLOCTPRQSGEYVYRAATVEAESGSIITLOVLATPGDLSCLVFKHSSLGQCPHD 120
Qy 120 LQNRGVSMVILKMTTQAGEYLLFIQSEATNYTLFTVSIINTLYTLRRPYFRMENO 179
Db 121 LQNRGIVSMAILNVTETQAGEYLLHIQSEBRANYTLFTVNVDQTOLYVLRPYFRMENO 180
Qy 180 DALVCISSEVPPIVWVLCDSQESCKEESPAVVKKEKVLHELFGTDIRCARNELGR 239
Db 181 DALLCISGEVPEPTVWVLCSSHRESCKEESPAVVKKEKVLHELFGTDIRCARNALGR 240
Qy 240 ECTRLFTIDLNOTPOTTLPOLFLKVGEPLEWIRCKAVHVNHGFLTWELNKALEGNVEE 299
Db 241 ECTKLFTIDLNOAPOSTLPOLFLKVGEPLEWIRCKAIVHVNHGFLTWELDKALEGSYFE 300
Qy 300 MSTYSTNRTMIRILFAPVSSVARNDTGYTCSSSRHPSQSALVTIVGKGFINATNSSEY 359
Db 301 MSTYSTNRTMIRILFAPVSSVGRNDTGYTCSSSRHPSQSALVTILEKGFINATNSQEBY 360
Qy 360 EIDQVEEFCFSVRFKAYPOIRCTWTFPSRKSFFCEOKGLDNGYSISKFNHKKHOPGEYIFH 419
Db 361 EIDPYEKEFCFSVRFKAYPRIRCTWTFPSQSAPFCEORGLDGYYSISKFDHKNKPGYIFY 420
Qy 420 AENDDAQFTKMTLNRKPOVLAEASASQSCFSDGYPLPSWTWKKCSKDPNCTEET 479
Db 421 AENDDAQFTKMTLNRKPOVLANASASQSCSDGYPLPSWTWKKCSKDPNCTEET 480
Qy 480 EGVNWKANRKYFGOWSSSTLNMEAIKGFVLKCCAYNSLGTSCETILNSPGPPFFQ 539
Db 481 EGVNWKANRKYFGOWSSSTLNMEAGLVLKCCAYNSMGTSCTETILNSPGPPFFQ 540
Qy 540 DNISFYATIGVCLLFTIVLTLIHKYKQFYESQLOMVQVGTSSDNEYFYVDREYEX 599
Db 541 DNISFYATIGVCLLFTIVLTLIHKYKQFYESQLOMVQVGTSSDNEYFYVDREYEX 600
Qy 600 DLKWFPPRENLEFGKVLGSGAGKVMNATAYGISTGVSIOVAVKMLKEKADSSEREA 659
Db 601 DLKWFPPRENLEFGKVLGSGAGKVMNATAYGISTGVSIOVAVKMLKEKADSSEREA 660
Qy 660 SELKMTQLGSHENIVNLGACTLSGPVILIPEYCCYGDLLNLYLRKREKFRHTWTEIFK 719
Db 661 SELKMTQLGSHENIVNLGACTLSGPVILIPEYCCYGDLLNLYLRKREKFRHTWTEIFK 720
Qy 720 ENHFSFYPTFQSHNPSMPSREVQIHPDSDOI SGLHNSFHSDEIEYENQRL--EE 777
Db 721 ENHFSFYPTFQSHNPSMPSREVQIHPDSDOI SGLHNSFHSDEIEYENQRLAE 780
Qy 778 EDNLVLTPEDDLCPAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDPGLARDIM 837
Db 781 EDNLVLTPEDDLCPAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDPGLARDIL 840
Qy 838 SDSNYVVRGNARILPVKWMAPESIFEGITYIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 897
Db 841 SDSNYVVRGNARILPVKWMAPESIFEGITYIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900
Qy 898 FYKLIQNGFKMPOPPYATBEIYIMOSCWAFPSRKRPSFNLTSFLGCOLADAEAEYQN 957
Db 901 FYKLIQNGFKMPOPPYATBEIYIMOSCWAFPSRKRPSFNLTSFLGCOLADAEAE 957
Qy 958 VDRGVSECPHTYQNRPPFRSREMDLGLLSPOAQVE 991
Db 958 ----IRTSIHLPKQAAPQORG-GLRAQSPORQVK 986

```

RESULT 5

```

US-10-192-867-2
; Sequence 2, Application US/10192867
; Publication No. US20030084466A1
; GENERAL INFORMATION:
; APPLICANT: BLUME-JENSEN, Peter
; APPLICANT: HUNTER, Tony
; TITLE OF INVENTION: SYSTEM AND METHOD FOR CONTROLLING MALE FERTILITY
; FILE REFERENCE: SALKINS.002C1

```

```

; CURRENT APPLICATION NUMBER: US/10/192,867
; CURRENT FILING DATE: 2002-07-08
; PRIOR APPLICATION NUMBER: US 60/175,625
; 2000-01-11
; PRIOR APPLICATION NUMBER: PCT/US01/00573
; 2001-01-08
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 975
; TYPE: PRT
; ORGANISM: MUS MUSCULUS
US-10-192-867-2

Query Match 24.0%; Score 1266; DB 15; Length 975;
Best Local Similarity 33.2%; Pred. No. 2.8e-94;
Matches 336; Conservative 173; Mismatches 330; Indels 172; Gaps 41;

Qy 35 CYLINHNKNDSSVGKSSSYPMVSESPEDLGCLARPOSSGTVVEAAAIVEVDVDSASITLQVL 94
Db 12 CVLLVLLRQGTATSPASPGPSPPP-----SIHPAQ-----ELIVEAGDTL--- 54
Qy 95 VDAPGNISCL-----WVFKHSSLNCQPHFDLQNRGVSMVILKMTTQAGEYLLFTQSE 148
Db 55 -----SLTICIDPDFVRWTFK-----TYFN-----EMVENKKNEMIQ-EKAE 89
Qy 149 ATNYTLFTVSIINTL-----LYTLRRPYFRMENQDALVCISESVPEPIVE 195
Db 90 ATR-TGTVTCNSNGLTSSIVFVRDPAPKFLVGLFLFGK-EDSDALVRCLPTDPQ-VSN 146
Qy 196 WVLCDQSQESCKEESPAV-----VKKEEKVLHELFGTDIRCARNELGRCTRL--- 244
Db 147 YSLIECDGKSLPTDLTFVFNPKAGITIKVKKAYRHLCL--VRCNAQ-----RDGTWLHSD 199
Qy 245 -FTIDLNOT-----POTTLPOL--FLKVGRLPLTRCKAVHVNHGFLTW-----E 286
Db 200 KTLTKVREAIKAIIPVSVETSHLLKGGDTFTVVTCTIKDVSTSVNMLKXNPOQHIQA 259
Qy 287 LENKALEEGNPEMSTYSTNRTMIRILFAPVSSVARNDTGYTCSSSKHPSQSALVT--- 343
Db 260 VKHNSWHRGDF---NYERQETLT-----ISSARVDDSGVPMCYANNTFGSANVTTLK 309
Qy 344 IVKGKFINATN--SSEDIQYEEFCFSVRFKAYPO-IRCTWTFPSRKSFFCEOKGLD--- 398
Db 310 VVEKGFINISPVKNTTFTVTDGENVDLVVEYAYPKPEHQWIMYMRN--SANKKDYVK 367
Qy 399 -NGYSISKFCNKH-----QPGEYI FEAENDDAQFTKMTLNRKPOVLA--EASAS 448
Db 368 SONKSNIRYVNOQLRLTRLKGTGGTYTFLVSNSDASAVTFNVYVNTKPBILTYDRLING 427
Qy 449 QASCFSDGYPLPSWTWKKCSKDPNCTEITE-GVMNRKANRKYFGOWSSSTLNMEAI 507
Db 428 MLQCVAGEPETIDWYFCTGAEQRCCTTPVSPVDVQVQNVSVSPFGKLVQSSIDSSVFR 487
Qy 508 KGFLVKCCAYNSLGTSCETILNSPGPFPF-----IQDN-----ISFYATIGVCLLPIV 557
Db 488 HNGTVECKASNDVGKS--SAFFN-----FAFKEQIQAHLTFTPLLLIGFVVAAGANGIIVM 541
Qy 558 LTLILCHKYKQFYESQLOMV-QVGTSSDNEYFYVDREYEDLKEWEPRENLEFGKVL 616
Db 542 LT-----YKYLQKPMYEVQWKVVEEING---NNYVYIDPTQLPYDHWKPEPRNLSFGKTL 594
Qy 617 GSGAFGKVMNATAYGISTGVSIOVAVKMLKEKADSSEREAALMSLKMMTQLGSHENIVN 676
Db 595 GAGAFGKVMNATAYGLIKSDAAMTAVAKMLKPSAHLTEREALMSLKVLSYLNHNHIVN 654
Qy 677 LLGACTLSGPYILPEYCCYGDLLNLYLRKREKFP-----HRTWTBIFKEHNFSPYPTFQ 730
Db 655 LLGACTVGGPTLVI TEYCCYGDLLNLYLRKREKFSIFSKQEEQAEAAALYKULLHSTBPSCD 714
Qy 731 SHPNSSM-----PGSREVOIHPDSDOI SGLHNSFHSDEIEYENQKLEEDNLVLTPE 787
Db 715 S-SNEYMDMKPGVSYV-VPTKDKRSARIDSY-----IERDVTPAIMDEDEL-ALDLD 766

```



Db 107 IYVVRDPKALFLVDRSLYK-EDNLTVRCLTDP-VTNYSKGCQCKPLKDLRPI 164  
QY 210 ---SNAVKEEVLHELFGTIRCCARNELEGTRELTIDL----NOTPQTTLPO--L 260  
Db 165 DPKAGIMIKSVKRAVHRLC---LHCSVDQEGKSVLSEKFIKVRPAKAVPVVSVKASY 221  
QY 261 FLKVGEPMLWIRKAVHVNHGFLTWELN---KALEBGNFYFEMSTYSTNRTMIRILFAPV 317  
Db 222 LLREGEFTVCTIKDVSSSVSTWKRNSQTKLEKYNWHGDPNYERQAT----LTI 277  
QY 318 SSVARNDTGYTCSKHPQSALVT---IVGKGFNA-----TNSSEDEYEDQYE 365  
Db 278 SSARVNDGVMFCYANNTFGSANVTTLLEVVDKGFNIPFMINITTVFVNDGENVDL---- 333  
QY 366 ECFVSVPKAYQ--TRCTWTFBRKSP--PCBOKGLDNGVSIKFCNKHK-----QPGE 415  
Db 334 ----IVEEAPFKPEHQIYNNRTFTDKWEDPKSESNIRYVSELHLRLKGTGEGT 389  
QY 416 YIFHAENDDAQFTKFTLIRKPOVLA--EASASQSCFSGYPLPSPWTKKCDKSPN 473  
Db 390 YTFVLSNDSVNAIAFNVVNTKPELITDRLVNGMLQCVAAAGFPEPTIDWTFPCGTEOR 449  
QY 474 CTEEITE-GVMNRKANRVFGQWSSSTLNMSEAIGFLVKCCAYNSLGTSCETILLNSP 532  
Db 450 CSASVLPVDVQTLNSSGPPFGKLVQSSIDSASFHNGTVECKAYNDVGKT--SAYFN-- 505  
QY 533 GPPPIQDN-----ISFYATIGVCLLPIVVLTLLIHKYKQKQRYESQLOM 578  
Db 506 --FAFKGNKEQIHPTLFTPLLIGFVIVAGMCIIVMLT----YKIQKPMYEQWKV 559  
QY 579 V-OVTGSSDNEFYVDYFREYEDLKEPRENLERKVLGSGAFKGMNATAYGSKTGV 637  
Db 560 VEEING---NNVYIDPTQLPYDHWEPFRNLSFGKTLGAGAFKQVEATAYGLIKSDA 616  
QY 638 SIQVAVKMLKEKADSESEALMSLKMMTQLGSHENIVNLLGACTLSGPIYLIPEYCCVG 697  
Db 617 AMTAVKMLKPSAHLTEREALMSLKVLKSYLGHNHNVNLLGACTIGGTLVITYCCVG 676  
QY 698 DLLNVLRSKREKPHRTWTE-----IFKEHNPFSFYPTFQSHPNSM--PGSREVOIHDP 748  
Db 677 DLLNPLRRKRDSFICSKQEDHAEALYKNLLHSKSSCSDSTNEYMDMKPGSVYV-VPTK 735  
QY 749 SDQISGLHNSHSEDELEYENQKLEEBEDLNVLTPELDFAYQVAKGMEFLFKSCV 808  
Db 736 ADKRSVRIGSY-----IERDVTIAMEDDEL-ALDLEDLLSFQVAKGMAFLASKNCI 789  
QY 809 HRDLAARNVLVTHGKVKICDPLGLARDIMSDSNVYVVRGNARLPVKWMAPESLFEGITYIK 868  
Db 790 HRDLAARNILLTHGRTIKICDPLGLARDIKNDSNVYVVGKGNARLPVKWMAPESLFNCVITYE 849  
QY 869 SDVMSYGILLWEIFSLGPNVPGIPVDANFYKLIQNGFKMDOPFYATEEIIYIMQSCWAF 928  
Db 850 SDVMSYGIFLWELFSLGSPYGPMPVDSFKYKMKIEGFRMLSPHAPAEYDIMKTCDWA 909  
QY 929 DSRKPSFNLTSFLGCOLADAEAMYQNV 958  
Db 910 DPLKREPTKQIVOLIEKQISESTNHIYSNL 939

## RESULT 8

US-09-944-807-10  
; Sequence 10, Application US/09944807  
; Patent No. US2002011949A1  
; GENERAL INFORMATION:  
; APPLICANT: Boehringer Ingelheim Pharma KG  
; TITLE OF INVENTION: Method for identifying substances which positively  
; TITLE OF INVENTION: Influence inflammatory conditions of chronic  
; TITLE OF INVENTION: inflammatory airway diseases  
; FILE REFERENCE: 082 00n  
; CURRENT APPLICATION NUMBER: US/09/944, 807  
; CURRENT FILING DATE: 2001-08-31  
; PRIOR APPLICATION NUMBER: UK 0021484.1

; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 972  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-944-807-10

Query Match 23.1%; Score 1216.5; DB IO; Length 972;  
Best Local Similarity 32.8%; Pred. No. 3.1e-90;  
Matches 325; Conservative 156; Mismatches 316; Indels 195; Gaps 33;

QY 75 VYEAARVVDV--SASITLOVLVDAPGNISCLWVFKHSSLNCPHDPDLQNRGVVSMVILK 132  
Db 22 VIEPSVPELVKPGATVTLRCV---GNGSVEMDGPSP---PHWTLTSDGSSSILSTN 72  
QY 133 MTE-----TOAGEYL-----LFTQSEATNVTLFTVSIIRNTLLYTLRRPYPRKM 176  
Db 73 NATPONTGYRCTEFGDPLGSSAAHLVYKDPARPNVL---AQEVVVF----- 118  
QY 177 ENQDALV-CISESVPEPIVEWVLCDSQGESCKEESPAVVKKEK-----VLHE--- 223  
Db 119 EDQDALLPCL--LTPDVL-----AGSVLRVRGRPLMRHTNYSFSFPMHGFTIHRAPF 169  
QY 224 LFGTDIRCCARNELEGTRELTIDLNOTPQTLP-----QLFKVGEPLMIRCK 273  
Db 170 IQSQDYQCSALMG-GR---KVMISIRLKVQKVIQPPALTLVPAELVRIRGEAAQIVCS 225  
QY 274 AVVNHGFLTWELNKALEBGNFYFEMSTYSTNRTMIRILFAPVSSVARNDTGYTCSSS 333  
Db 226 ASSVDVNFDFLQHNNTKLA---IPQSDFNHNRYO-KVLTNLNDQVDPQHAGNYSVCAS 281  
QY 334 ----KHPSQSALVTIVGKGFNATNSSD---YEIDQYEEFCFVSFRKAYPOIR-CTWTF 385  
Db 282 NVQGHK-STSMFRVVESAYLNL--SSQNLIQEVTVGEGNLKVMVEAYPGLOGFNWTY 338  
QY 386 SRKSPCEQKGLDNGYSISKFCNKHQP-----GEYIP 418  
Db 339 -----LGPPFDHQPEPKLANATTKTDTYRHTFTLSLPLKLPSEAGRYSF 381  
QY 419 HAENDDAQFTKFTLIRKPOVLAESASQAS-----CFSDGYPLPSPWTKKCDKSPNC 474  
Db 382 LARNPGWRALTFELTLRYPEVSVVITFINGSGTLLCAASGYPOPNVTWLCQSGHTDRC 441  
QY 475 TEEITEGVN---RKANRVFGQWSSSTLNMSEAIGFLVKCCAYNSLGTSCETILLN 530  
Db 442 DEAQVLQWDDPYPEVLSQEPFKVTVQSLLTVETLEHNTVECAHNSVSGSWAFIFI 501  
QY 531 SPGPPPPFIQDNISFYATIGVC---LLFTIVLTLLICHYKQKQRYESQLOMVOVTGSSDN 587  
Db 502 SAGARTHPDDEFLTPVVVACWSIMALLLLLLLYKQKQRYESQLOMVOVTGSSDN 559  
QY 588 EYFYVDREYDYDLKWEPPRENLEPKVLGSGAFKGMNATAYGSKTGVSIQVAVKMLK 647  
Db 560 SYTFIDPTQLPYNEKWEPPRNLOFGKTLGAGAFKGVVEATAFGLGKEDAVLKVAVKMLK 619  
QY 648 EKADSSREALMSLKMMTQLGSHENIVNLLGACTLSGPIYLIPEYCCVGDLLNVLRSKR 707  
Db 620 STAHADEKALMSLKMMTQLGSHENIVNLLGACTHGGFVLVITYCCVGDLLNVLRSKR 679  
QY 708 E-----KFHRTWTEIFKEHNPFSFYPTFQSHPNSSMPGSRREVQIHP 747  
Db 680 EAMLGPSLSPGQDPGEGVDYKNIHLEKKYVRDSCFS-----SQGVDTIV 724  
QY 748 DSDQISGLHNSHSEDELEYENQKLEEBEDLNVLTPELDFAYQVAKGMEFLFKSC 807  
Db 725 EMRPVSTSSNDSF-SEQDLKEDGRPLE-----LRDLLHSSQVQAGWAFASKNC 774  
QY 808 VHRDLAARNVLVTHGKVKICDPLGLARDIMSDSNVYVVRGNARLPVKWMAPESLFEGITYI 867  
Db 775 IHRDVAARNVLLTNGHAKIGDFGLARDIMSDSNVYVVRGNARLPVKWMAPESLFPCVTV 834

QY 868 KSDVWSYGILLWEISLGNVPYGPVDFANFYKLIQNGFKMDOPFYATEEIIYIMOSCA 927  
DB 835 QSDVWSYGILLWEISLGNVPYGPVDFANFYKLIQNGFKMDOPFYATEEIIYIMOSCA 894  
QY 928 FDSRRKRPFPNLTSLFGCOL-ADAEAMYQNV 958  
DB 895 LEPTHPFTFOQICSFLQEQAOEDRERDYTNL 926

RESULT 9  
US-09-961-403-4  
; Sequence 4, Application US/09961403  
; Publication No. US2003007589A1  
; GENERAL INFORMATION:  
; APPLICANT: HE-STUMPP, HOLGER  
; APPLICANT: HAENDLER, BERNARD  
; APPLICANT: KRAETZSCHMAR, JOERN  
; APPLICANT: KREFT, BERTHOLT  
; APPLICANT: WINTERHAGER, ELKE  
; APPLICANT: REGIDOR, PEDRO  
; APPLICANT: SCOTTI, SIMONE  
; TITLE OF INVENTION: METHOD FOR IN VITRO DIAGNOSIS OF ENDOMETRIOSIS  
; FILE REFERENCE: SCH-1789  
; CURRENT APPLICATION NUMBER: US/09/961,403  
; CURRENT FILING DATE: 2001-09-25  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 1088  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-961-403-4

Query Match 22.1%; Score 1166.5; DB 11; Length 1088;  
Best Local Similarity 30.4%; Pred. No. 4.4e-86;  
Matches 336; Conservative 165; Mismatches 365; Indels 241; Gaps 36;

QY 12 PLLVVSAMIFG---TITNQDLPLVI-----KCVLINHKNDSSVGKSS---SYPMVSES 59  
DB 6 PAFLVLGCLLTGLSLILCQLSLPLPNEKENVQVQLNSSFSLRCFGESEVSWQPMSEEE 65  
QY 60 PEDLCALRPQSG---TYEAAAEVVDVSAITQLVLDAPGNISCLVFKHSHLNCOP 116  
DB 66 SSDVEIRNEENNSGLFVTVLEVSS-----ASAAHT-----GLTYCYNHTQTEEN--- 110  
QY 117 HFDLQNRGV-----VSMVILKMTETQAGVYLLFIOSEATNYTILFTVSIRNTLYTL 168  
DB 111 --ELEGRIIYVPPDVAFLPGLMT-----DYLIVE----- 141  
QY 169 RPYFRKMNODALVCISESVPEPIVWVLCDSQG-----ESCKE 208  
DB 142 -----DDSAIIPCRITTDPETPV--TLHNSGVVPASVDSRQGFNGFTVGPVIC-- 189  
QY 209 ESPAVVKEEKVHLFGTDISCCARNELGRECTRLFTIDLNQPTTLPQLFLKVGEP 268  
DB 190 EATVGRKFPQTFIPFNVALK-----ATSELDLEMEALKTV-----YKSGETI 231  
QY 269 WIRCKAVHNVHFGLTW-----ELENKA-----LEEGNVFEMSTYTNRTMIRILFAP-VSSV 320  
DB 232 VUTC-AVFNEVVDVLTWTPGVGKGTMLSE-----IKVPSIKLVITLTVPBA 280  
QY 321 ARNDTGYTCSK-----HPQSALVTIVGKFINATNSSEDEY-IDQYEEFCFSVRPK 374  
DB 281 TVKDSGDVECAARQATREVKEMKVTISVHENGFIKPTFPQLEAVNLHEVKHFVVR 340  
QY 375 AYPQIRCTWTFRRKFPCEPKGLDNGYSIS-----KFCNHKIQ-PGEYIF 418  
DB 341 AYPPIRISMLKNLTL-----IENLTETDVEKIQEIRYSKULIRAKEEDSGHYTI 394  
QY 419 HAENDAOPTKMTNIRKPKOVL-----AEASASQACFSDGYPLPSWTWKKCSK 471  
DB 395 VAQNEDAVKSYTFELLTQVPSILDLVDHHSSTGGQTVRCAEGTPLPDIEWMICKO-I 453

QY 472 PNCBTEITEGVNWKANRKFVQWVSS-----STLNMSEAIKGLYKCCAY 517  
DB 454 KCCNNETS---WTLANN-----VSNITTEIHSRDRSTVEGRVTFKAVEETIAVRCLAK 504  
QY 518 NSLGTSCETILLNSGPPFPFIQDNISFVATICVCLLFIVLWLLIICHKVKKQFRYESOLQ 577  
DB 505 NULGARENRELKVA---PTLESELTVAAAVLVIVIIISLIVLVVWKQKREIWR 560  
QY 578 MVQVTGSSDNEYFYVDREYEDLKWEPFRENLEFGKVLGSCAFGKVMNATAYGISKTV 637  
DB 561 VIESISPDGHEIYVDPMLPYDSRWEFPRDGLVGLRGVSGAFGKVEGTAYGLRSOP 620  
QY 638 SIQAVAKMLKESKADSSEREALMSKMTQLGSHENIYNLLGACTLSGDIYILIFEYCCY 697  
DB 621 VMKAVAKMLKPTARSSEKQALMSKIMTHLGPLHNLVNLGACTKSGDIYIITEYCFY 680  
QY 698 DLLNLYRSKREKF-----HRTWTEIFKEHNSFYPTFQSHPN 734  
DB 681 DLVNLHKNRDSFLSHHPKPKKELDIPLGNPADESTSYVILSFENNDDYMDMKQADTT 740  
QY 735 SSMGSRVQIHPDSQISGLHGN--SFHSEDEIEYENQKRLSEEDLNLVTFEDLLCPA 792  
DB 741 QVVPMLERKEYSKYSIQISLYDRPASVYKKSMLDSE-VKNLLSDDNSEGLTLLDLSFT 799  
QY 793 YQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKIKCDFGLARDIMSDSYVVRGNARLP 852  
DB 800 YQVARGMEFLASKNCVHRDLAARNVLLAQGKIVKICDFGLARDIMHDSNYVSGSTFLPV 859  
QY 853 KWMAPESLFEGITIKSDVMSYGILLWEIFSGVNPYGPVDFANFYKLIQNGFKMDOP 912  
DB 860 KWMAPESLFDMLYTTLSDVMSYGILLWEIFSGVNPYGPVDFANFYKLIQNGFKMDOP 919  
QY 913 YATEEIIYIMOSCAWAFDSKRKPSFNLTSFLGCQLADAEAMYQNV--DGRVSECPH--- 967  
DB 920 HATSEVYIMVKWNSSEKPSFYHLSEIVENLLPGQYKSYEKIHLDFLAKSDHPAVAR 979  
QY 968 -----TYQNRPRFSPREMDLGL 983  
DB 980 MVDSNDAYIGVTYKNEEDKLDKDWEGGL 1006

RESULT 10  
US-09-769-987-2  
; Sequence 2, Application US/09769987  
; Patent No. US20020055129A1  
; GENERAL INFORMATION:  
; APPLICANT: Matsui, Toshimitsu  
; APPLICANT: Aaronson, Stuart A.  
; APPLICANT: Pierce, Jacalyn H.  
; TITLE OF INVENTION: Antibodies for the Alpha Platelet-Derived Growth Factor Receptor;  
; FILE REFERENCE: 14014.0266U2  
; CURRENT APPLICATION NUMBER: US/09/769,987  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 08/460,656  
; PRIOR FILING DATE: 1995-06-02  
; PRIOR APPLICATION NUMBER: US 08/439,095  
; PRIOR FILING DATE: 1995-05-11  
; PRIOR APPLICATION NUMBER: US 07/915,884  
; PRIOR FILING DATE: 1992-07-20  
; PRIOR APPLICATION NUMBER: US 07/308,282  
; PRIOR FILING DATE: 1989-02-09  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 1089  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence; No. US20020055129A1 =  
; OTHER INFORMATION: synthetic construct  
US-09-769-987-2

Query Match 22.1%; Score 1166; DB 9; Length 1089;







Kindsvogel Ph.D., Wayne R.  
 TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS  
 AND BIOLOGICALLY ACTIVE DIMERIZED POLYPEPTIDE  
 FUSIONS  
 NUMBER OF SEQUENCES: 36  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Seed and Berry  
 STREET: 6300 Columbia Center  
 CITY: Seattle  
 STATE: WA  
 COUNTRY: USA  
 ZIP: 98104-7092  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.24  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/955,363  
 FILING DATE: 18-Sep-2001  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/634,510  
 FILING DATE: <Unknown>  
 APPLICATION NUMBER: US 07/347,291  
 FILING DATE: 02-MAY-1989  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Maki J.D., David J.  
 REGISTRATION NUMBER: 31,392  
 REFERENCE/DOCKET NUMBER: 990008.446C3  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 206-622-4900  
 TELEFAX: 206-682-6031  
 TELEX: 3723836  
 INFORMATION FOR SEQ ID NO: 36:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1089 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 36:  
 US-09-955-363-36  
 Query Match 22.1%; Score 1166; DB 10; Length 1089;  
 Best Local Similarity 30.3%; Pred. No. 4.8e-86;  
 Matches 336; Conservative 165; Mismatches 365; Indels 242; Gaps 36;  
 QY 12 PLLVVFSAIFG---TITNQLPVI-----KCVLINHKNDSSVCKSS---SYPMVSES 59  
 DB 6 PAFLVGLCLTGLSLILCOLSLPSILPNEKENVQLNSFSLRFCGESEVSWOYPMSEEE 65  
 QY 60 PEDLGALRPOSSG---TYEAAVEVDVSAITLQVLVDAPGNISCLMVFVKHSSLNCP 116  
 DB 66 SSDVEIRNEENNSGLFVTVLEVS-----ASAHT-----GLYTCYNHQTSEN--- 110  
 QY 117 HPDLQNRGV-----VSMVLKWTETQAGYLLFIQSEARNYTLFTVSRINTLLYTL 168  
 DB 111 --ELEGRHIYIVPDPAFVPLGMT-----DYLVIIE----- 141  
 QY 169 RPYPRKMNQDALVCISEVPEPIVEWVLCDSQ-----ESCKE 208  
 DB 142 -----DDSAIIPCRITDPETV--TLHNSGVVPASVDSROGFNGFTTVGPYIC-- 189  
 QY 209 ESPAVVKEEKLHFGDIRCCARNELGRCTRLFTIDLNQTPQTLPLQFLKVGEP 268  
 DB 190 EATVKGKFIQIPFNVAALK-----ATSELDLEMEALKTV-----YKGETI 231  
 QY 269 WIRKAVHNVHGFGLTW-----ELENA-----LEEKNYFEMSTYSTNRTMIRILFAF-VSSV 320  
 DB 232 VVTC-AVENNEVVDLQWTPGVEKVGKITILEE-----IKVPSIKLYVTLTVPEA 280  
 QY 321 ARNDGTGYTCSSK-----HFSQSALVTIVGVGFNATNNSSEDEY-IDQYEFECFSVRPK 374

DB 281 TVKSDGYECAARQATREVKEMKQVTVISVHEKGFIEIKFTFSQLEAVNLHEVGHFVVEVR 340  
 QY 375 AYPQIRCTWTFPSRKPCCQKGLDNGYSIS-----KFCNHHKQ-PGEYIF 418  
 DB 341 AYPFPRISLWLNKNTL-----IENLTETDVEKIQEIRYSKLIIRAKEDSGHYTI 394  
 QY 419 HAENDDAOFTKMTLNIIRKQVPL-----AEASASQASCFSDDGYPLPSWTWKKCSDKS 471  
 DB 395 VAQNEADVKSFTFELLTQVPSSILDVDDHSGSTGGTVRCTAEGTPLPDIEWICKD-I 453  
 QY 472 PNCTEEITGVWNRKANRKFQGVSS-----STLNMSEAIKFLVKCCAY 517  
 DB 454 KCKNNETS---WTILANN-----VSNIIITEHSRDRSTVEGRVTFKVEETIAVRCLAK 504  
 QY 518 NSLGTSCETILLNSPGPPFIODNISFYATIGVCLLFIVVLTLILCHIKVKKQPRYSQLQ 577  
 DB 505 NLLGAENRELKIVA---PTLSELTAAAVLVLVIVLIVLIVLVIWQKQPRYIRWR 560  
 QY 578 MYQVTGSSDNEYFYVDREYEDLKWEPRENLEKGLVSGAGFKVMNATAYGISKTGV 637  
 DB 561 VIESISPDGHEIYYVDPOLPYDSRWEFPRDGLVLRVLGSGAFKVGEGTAYGLSRQP 620  
 QY 638 SIQVAVKMLKEKADSSEREALMSLKMMTQLGSHENIVNLLGACTLSGPTILIFEYCCYG 697  
 DB 621 VMKVAVKMLKPTARSSEKQALMSLKMTHLGPHLNIVNLLGACTSGPIIITEYCFYG 680  
 QY 698 DLLNLYLSRKREKF-----HRTWTEIFKEHNFSEFYPTFQSHPN 734  
 DB 681 DLVNYLHKNRDSFLSHHPEKPKKELDIFGLNPADESTRSYVILSPENNGDYMDMKQADTT 740  
 QY 735 SSMPGSRVQHPDSDQISGLHGN---SPHSEDEIEYENOKRLEEEDLNVLPEDLLCPA 792  
 DB 741 QYVPMLEKREKVSYSIDIQRSLYDRPASYYKKSMLEDS-VKNLLSDNSEGLTLLDLSFT 799  
 QY 793 YQVAKGMEFLPKSCVHRDLAARNVLVTHGKVKVVICDFGLARDIMSDSNVVRGNARLPV 852  
 DB 800 YQVARGMEFLASKNCVHRDLAARNVLLAQGKIVKICDFGLARDIMHDSNVYKSGTFLPV 859  
 QY 853 KWMAPELFEGIYTIKSDVMSYGILLWEIFSLGVNVPYGPIDVNFYKLIQNGFKMDQPF 912  
 DB 860 KWMAPELFDNLYTTLSDVMSYGILLWEIFSLGGTYPYGMVVDSTFYNKIKSGYRMAKPD 919  
 QY 913 YATEIYIIMQSCWAFDSRKPSPFNLTSLFQCQLADAEEMYNQV--DGRVSECPH--- 967  
 DB 920 HATSEVYIIMVKNSEPEKRPSPFYHLSIEVENLLPGQYKSYEKIHLDFLAKSDHPAVAR 979  
 QY 968 -----TYQNRPRPFSREMDLGL 983  
 DB 980 MRVSDNAYIGVTYKNEEDKLDKDWEGGL 1007  
 RESULT 14  
 US-09-866-510-10  
 ; Sequence 10, Application US/09866510  
 ; Patent No. US20020111304A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: KAZLAUSKAS, ANDRIUS  
 ; APPLICANT: IKUNO, YASUSHI  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES  
 ; FILE REFERENCE: ERM-104.01  
 ; CURRENT APPLICATION NUMBER: US/09/866,510  
 ; CURRENT FILING DATE: 2001-05-25  
 ; PRIOR APPLICATION NUMBER: 60/250,747  
 ; PRIOR FILING DATE: 2000-12-01  
 ; PRIOR APPLICATION NUMBER: 60/289,103  
 ; PRIOR FILING DATE: 2001-05-07  
 ; NUMBER OF SEQ ID NOS: 33  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 10  
 ; LENGTH: 1089  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-866-510-10

```
Query Match      22.1%; Score 1163; DB 10; Length 1089;
Best Local Similarity 30.2%; Pred. No. 8.5e-86;
Matches 335; Conservative 166; Mismatches 365; Indels 242; Gaps 36;

QY 12 PLLVFSAMIFG---TITNODLPVI-----KCVLNHNKNDSSVCKSS---SYPMVSES 59
DB 6 PAFLVLGCLLTGLSLILCQLSLPILPNEKVKVQLNSSFRLCFGESESVWQYPMSEEE 65

QY 60 PEDLGCALPQSSG---TVVEAAAVEVDVSASITLQVLVDAPGNISCLMWFKHSSLNCQP 116
DB 66 SSDVEIRNEENSGLFVTVLEVSS---ASAHT-----GLTCYNNHTQTEEN--- 110

QY 117 HFDLQNRGV-----VSMVILKMTETOAGEYLLFIQSEATNYTLFTVSRINTLLYTL 168
DB 111 --ELEGRIHYIVPDPAFVPLGMT-----DYLIVE----- 141

QY 169 RRPYFRKMNQDALVCISSEVPPIVWVLCDSQG-----ESCKE 208
DB 142 -----DDSAIIPCRITTDPTPV--TLHNSEGVVPASYSRQGFNGFTVGPYIC-- 189

QY 209 ESPAVVVKKEKVLHFLGFTDIRCCARNELGRECTRLFTIDLNOTPTTLPLQFLKVGCEPL 268
DB 190 EATVKGKQTFIPFNVAALK---ATSELDLEMEALKTV-----YKSGEII 231

QY 269 WIRKCAVHNHGFGLTW---ELENKA---LEEGNYFEMSTYSTNRTMIRILPAP--VSSV 320
DB 232 VVTC-AVFNEVVLDQWTPGVEKVGKGTWLEB-----IKVPSIKLVYTLTVPFA 280

QY 321 ARNDGTYYTCSKSK-----HPSQSALVTIVGKGFINATNSSEDEY-IDQVEEFCFSVRFK 374
DB 281 TVKDSGDYECARQAATREVKEMKKTIVSHEKGFIEIKPTFSQLEAVNLHEVKHFVVEVR 340

QY 375 AYPQIRCTWTFRSKSPCEQKGLDNGYSIS-----KFCNKHQ--PGEYIF 418
DB 341 AYPPPPISLKNLTL-----IENLTETTDVEKIQEIRYRSKLLIRAKESDSGHYII 394

QY 419 HAENDDAQTKMFTLNIIRKPOVL-----AEASASQSCFSDGVPPLSWTWKCSKDS 471
DB 395 VAQNEADAVKSYTFELLTOVPPSSILDLVDDHGGSTGGTVCRTAEGTPLDIEWMICKD-I 453

QY 472 PNCTEITGVWNRKANRKFVGOWSS-----STLNMSIAIKGLVKCCAY 517

Query Match      22.0%; Score 1162; DB 10; Length 1089;
Best Local Similarity 30.2%; Pred. No. 1e-85;
Matches 335; Conservative 166; Mismatches 365; Indels 242; Gaps 36;

QY 12 PLLVFSAMIFG---TITNODLPVI-----KCVLNHNKNDSSVCKSS---SYPMVSES 59
DB 6 PAFLVLGCLLTGLSLILCQLSLPILPNEKVKVQLNSSFRLCFGESESVWQYPMSEEE 65

QY 60 PEDLGCALPQSSG---TVVEAAAVEVDVSASITLQVLVDAPGNISCLMWFKHSSLNCQP 116
DB 66 SSDVEIRNEENSGLFVTVLEVSS---ASAHT-----GLTCYNNHTQTEEN--- 110

QY 117 HFDLQNRGV-----VSMVILKMTETOAGEYLLFIQSEATNYTLFTVSRINTLLYTL 168
DB 111 --ELEGRIHYIVPDPAFVPLGMT-----DYLIVE----- 141

QY 169 RRPYFRKMNQDALVCISSEVPPIVWVLCDSQG-----ESCKE 208
DB 142 -----DDSAIIPCRITTDPTPV--TLHNSEGVVPASYSRQGFNGFTVGPYIC-- 189

QY 209 ESPAVVVKKEKVLHFLGFTDIRCCARNELGRECTRLFTIDLNOTPTTLPLQFLKVGCEPL 268
DB 190 EATVKGKQTFIPFNVAALK---ATSELDLEMEALKTV-----YKSGEII 231

QY 269 WIRKCAVHNHGFGLTW---ELENKA---LEEGNYFEMSTYSTNRTMIRILPAP--VSSV 320
DB 232 VVTC-AVFNEVVLDQWTPGVEKVGKGTWLEB-----IKVPSIKLVYTLTVPFA 280

QY 321 ARNDGTYYTCSKSK-----HPSQSALVTIVGKGFINATNSSEDEY-IDQVEEFCFSVRFK 374
DB 281 TVKDSGDYECARQAATREVKEMKKTIVSHEKGFIEIKPTFSQLEAVNLHEVKHFVVEVR 340

QY 375 AYPQIRCTWTFRSKSPCEQKGLDNGYSIS-----KFCNKHQ--PGEYIF 418
DB 341 AYPPPPISLKNLTL-----IENLTETTDVEKIQEIRYRSKLLIRAKESDSGHYII 394

QY 419 HAENDDAQTKMFTLNIIRKPOVL-----AEASASQSCFSDGVPPLSWTWKCSKDS 471
DB 395 VAQNEADAVKSYTFELLTOVPPSSILDLVDDHGGSTGGTVCRTAEGTPLDIEWMICKD-I 453

QY 472 PNCTEITGVWNRKANRKFVGOWSS-----STLNMSIAIKGLVKCCAY 517
```

Db 454 KCCNNETS---WTILANN-----VSNITEIHSRDRSTVEGRVTEFAKVEETIAVRCLAK 504  
QY 518 NSLGTSCETILLNSPGPFPIQDNISFYATIGVCLLPIVVLTLILCHYKQPRYESQLQ 577  
Db 505 NULGAENRELKVA-----PTLRSELVAAAVLLVLIIVISLIVLVIWVKQPRYEIRWR 560  
QY 578 MVQVTGSSDNEFYVDFREYEDLKWEFPRENLEFGKVLGSGAFGKVMNATAYGISKTGV 637  
Db 561 VIESISPDGHEIYVDPQMLPYDRWKFRDGLVLRVLGSGAFGVVEGTAYGLRSQP 620  
QY 638 SIQVAVKMLKEKADSSERALMSLKMOTQLGSHENIVNLLGACTLGGPIYLIPEYCCYG 697  
Db 621 VMKAVKMLKPTARSSEKQALMSLKMTHLPHLNIIVNLLGACTKSGPIYIITEYCFYG 680  
QY 698 DLLNVLRSKREKF-----HRTWTEIFKEHNFSPYPTFQSHEN 734  
Db 681 DLVNYLHKNRDSFLSHHPEKPKKELDIFGLNPADESTRSYVILSPENNGDYMDMKQADTT 740  
QY 735 SSMPGSRVQIHPDSDQISGLHGN--SPHSEDEIEYENOKRLEEEEDLNLVTFEDLLCFA 792  
Db 741 QYVPMLEKEVSKYSDIQBSLYDRPASYYKKSMLDSE-VKNLLSDDNSEGLTLLDLSET 799  
QY 793 YQVAKGMEFLEPKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDSNYVVRGNARLFPV 852  
Db 800 YQVARGMEFLASKNCVHRDLAARNVLAQKIVKICDFGLARDIMHDSNYVSKGSTFLFPV 859  
QY 853 KWMAPESLPEGIYTIKSDVWSYGILLWEIFSLGVNPPGIPVDANPYKLIQNGFKMDQPF 912  
Db 860 KWMAPESIFDNLTYTLSDVWSYGILLWEIFSLGSGTYPGGMVDSFYNNKIKSGYRMAKPD 919  
QY 913 YATEEIIYIMQSCWAFDSRKPSFNLTSFLGCQLADAEAMYNQV--DGRVSECPH--- 967  
Db 920 HATSEVYIMVKWNSPEKPSFYHLSEIVENLLPGQYKKSXEKIHLDPLKSDHPAVAR 979  
QY 968 -----TYQNRPRPFSREMDLGL 983  
Db 980 MRVDSDNAYIGVTYKNEEDKLADWEGGL 1007

Search completed: August 26, 2003, 07:41:12  
Job time : 50.5229 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2003, 07:21:59 ; Search time 21.5108 Seconds  
(without alignments)  
1953.187 Million cell updates/sec

Title: US-09-919-408A-4  
Perfect score: 5274  
Sequence: 1 MPALARDAGTVPLLVFSAM.....PFSREMDLGLSPQAQVEDS 993

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5274	100.0	993	1 US-07-977-451-4	Sequence 4, Appli
2	5274	100.0	993	1 US-08-252-517-4	Sequence 4, Appli
3	5274	100.0	993	1 US-07-906-397A-4	Sequence 4, Appli
4	5274	100.0	993	1 US-08-601-891-4	Sequence 4, Appli
5	5274	100.0	993	2 US-09-021-324-4	Sequence 4, Appli
6	5274	100.0	993	5 PCT-US92-09893-4	Sequence 4, Appli
7	5274	100.0	1160	5 PCT-US92-05401-4	Sequence 4, Appli
8	5240	99.4	993	1 US-08-222-299-4	Sequence 4, Appli
9	5240	99.4	993	2 US-08-434-878-4	Sequence 4, Appli
10	5240	99.4	993	5 PCT-US95-03718-4	Sequence 4, Appli
11	5238	99.3	993	1 US-08-183-211-2	Sequence 2, Appli
12	5238	99.3	993	5 PCT-US95-00176A-2	Sequence 2, Appli
13	4533.5	86.0	1000	1 US-08-222-299-2	Sequence 2, Appli
14	4533.5	86.0	1000	2 US-08-434-878-2	Sequence 2, Appli
15	4533.5	86.0	1000	5 PCT-US95-03718-2	Sequence 2, Appli
16	4429.5	84.0	992	1 US-07-813-593-2	Sequence 2, Appli
17	4429.5	84.0	992	1 US-07-977-451-2	Sequence 2, Appli
18	4429.5	84.0	992	1 US-07-946-507-2	Sequence 2, Appli
19	4429.5	84.0	992	1 US-08-252-517-2	Sequence 2, Appli
20	4429.5	84.0	992	1 US-07-906-397A-2	Sequence 2, Appli
21	4429.5	84.0	992	1 US-08-601-891-2	Sequence 2, Appli
22	4429.5	84.0	992	2 US-09-021-324-2	Sequence 2, Appli
23	4429.5	84.0	992	5 PCT-US92-02750-2	Sequence 2, Appli
24	4429.5	84.0	992	5 PCT-US92-05401-2	Sequence 2, Appli
25	4429.5	84.0	992	5 PCT-US92-09893-2	Sequence 2, Appli
26	1654	31.4	481	3 PCT-US92-122-4	Sequence 4, Appli
27	1654	31.4	481	5 PCT-US93-06404-4	Sequence 4, Appli

28 1251 23.7 976 3 US-08-750-141A-1 Sequence 1, Appli  
29 1216.5 23.1 972 3 US-08-750-141A-2 Sequence 2, Appli  
30 1166 22.1 1089 1 US-08-180-195-36 Sequence 36, Appli  
31 1166 22.1 1089 1 US-08-168-917-4 Sequence 4, Appli  
32 1166 22.1 1089 1 US-08-477-329-36 Sequence 36, Appli  
33 1166 22.1 1089 2 US-08-475-458-36 Sequence 4, Appli  
34 1166 22.1 1089 2 US-08-460-510-4 Sequence 36, Appli  
35 1166 22.1 1089 3 US-08-460-490-4 Sequence 4, Appli  
36 1166 22.1 1089 3 US-08-980-400-36 Sequence 36, Appli  
37 1166 22.1 1089 3 US-08-462-728-2 Sequence 2, Appli  
38 1166 22.1 1089 3 US-09-583-459A-36 Sequence 36, Appli  
39 1166 22.1 1089 3 US-09-583-210-36 Sequence 36, Appli  
40 1166 22.1 1089 4 US-09-583-449A-36 Sequence 36, Appli  
41 1166 22.1 1089 4 US-09-435-059-36 Sequence 36, Appli  
42 1166 22.1 1089 4 US-08-461-917-2 Sequence 2, Appli  
43 1166 22.1 1089 4 US-08-464-436-2 Sequence 2, Appli  
44 1166 22.1 1089 5 PCT-US92-00730-4 Sequence 4, Appli  
45 1166 22.1 1089 5 PCT-US92-00862-4 Sequence 4, Appli

## ALIGNMENTS

RESULT 1  
US-07-977-451-4  
; Sequence 4, Application US/07977451  
; Patent No. 5270459  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ImClone Systems Incorporated  
; STREET: 180 Varick Street  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/07/977,451  
; FILING DATE: 19921119  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US UNASSIGNED  
; FILING DATE: 12-NOV-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/906,397  
; FILING DATE: 26-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US PCT/US92/05401  
; FILING DATE: 26-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: TW 81102961  
; FILING DATE: 15-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US PCT/US92/02750  
; FILING DATE: 02-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/813,593  
; FILING DATE: 24-DEC-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/793,065  
; FILING DATE: 15-NOV-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/728,913  
; FILING DATE: 28-JUN-1991  
; PRIOR APPLICATION DATA:



APPLICATION NUMBER: US 07/679,666  
 FILING DATE: 02-APR-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Feit, Irving N.  
 REGISTRATION NUMBER: 28,601  
 REFERENCE/DOCKET NUMBER: LEM-3-7P  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 212-645-1405  
 TELEFAX: 212-645-2054  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 993 amino acids  
 TYPE: AMINO ACID  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-07-977-451-4

Query Match 100.0%; Score 5274; DB 1; Length 993;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPVLLVFSAMIFGTITNQDLPIKCVLINHKNNDSSVGKSSSYPMVSESP 60  
 Db 1 MPALARDAGTVPVLLVFSAMIFGTITNQDLPIKCVLINHKNNDSSVGKSSSYPMVSESP 60  
 QY 61 EDLGCALRQSSGTVEAAAEVDVDSASITLQVLVDAPGNISCLWPKHSSLNCQPHFDL 120  
 Db 61 EDLGCALRQSSGTVEAAAEVDVDSASITLQVLVDAPGNISCLWPKHSSLNCQPHFDL 120  
 QY 121 QNRGVVSWILKMTETQAGYLLFIQSEATNYTLFTVSRNTLLYTLRPPYPRKMNQD 180  
 Db 121 QNRGVVSWILKMTETQAGYLLFIQSEATNYTLFTVSRNTLLYTLRPPYPRKMNQD 180  
 QY 181 ALVCISESPEPIVEVWVLCDSQESCKEESPAVVKEEVLHFLFGTDIRCCARNEIGRE 240  
 Db 181 ALVCISESPEPIVEVWVLCDSQESCKEESPAVVKEEVLHFLFGTDIRCCARNEIGRE 240  
 QY 241 CTRLFTIDLNQTTLPQLFLKVGEPWIRCAVHVNHGFGTLWELENKALSEGNYFEM 300  
 Db 241 CTRLFTIDLNQTTLPQLFLKVGEPWIRCAVHVNHGFGTLWELENKALSEGNYFEM 300  
 QY 301 STYSTNRMTIRILFAPVSSVARNDGTGYTSSSKHPSQSALVTIVGKFINATNSSEDEYE 360  
 Db 301 STYSTNRMTIRILFAPVSSVARNDGTGYTSSSKHPSQSALVTIVGKFINATNSSEDEYE 360  
 QY 361 IDQYEFCEFSVRKAPQIRCTWTFRSKPPCEQKGLDNGYSISKFCNKHQGEYIFHA 420  
 Db 361 IDQYEFCEFSVRKAPQIRCTWTFRSKPPCEQKGLDNGYSISKFCNKHQGEYIFHA 420  
 QY 421 ENDDAQFTKMTLNIRKPKQVLAESASQASCFSDGYPLPSWTWKCKSDKSPNCTEBITE 480  
 Db 421 ENDDAQFTKMTLNIRKPKQVLAESASQASCFSDGYPLPSWTWKCKSDKSPNCTEBITE 480  
 QY 481 GVNVRKANRKFVQGVSSSTLNSEATKGLVKCCAYNSLGTSCETILLNSPGPFPIQD 540  
 Db 481 GVNVRKANRKFVQGVSSSTLNSEATKGLVKCCAYNSLGTSCETILLNSPGPFPIQD 540  
 QY 541 NISFYATIGVCLLFIIVVTLILCHYKKQFRESQLOMVQVSSDNEYFVFPREYEYD 600  
 Db 541 NISFYATIGVCLLFIIVVTLILCHYKKQFRESQLOMVQVSSDNEYFVFPREYEYD 600  
 QY 601 LKWEFFRENLEFGKVLGSGAFGKVMNATAYGISKTGVSIOQVAVMLKEKADSSEREALMS 660  
 Db 601 LKWEFFRENLEFGKVLGSGAFGKVMNATAYGISKTGVSIOQVAVMLKEKADSSEREALMS 660  
 QY 661 ELKQMTOLGSHENIVNLGACTLSGPTLYLFEYCCYGLNLYLSKKEKPHRTWTETPFKE 720  
 Db 661 ELKQMTOLGSHENIVNLGACTLSGPTLYLFEYCCYGLNLYLSKKEKPHRTWTETPFKE 720  
 QY 721 HNFSPYFTFOSHNSMPGSGREVQIHFDSDQISGLHGNSFHSDEIEYENQKLEEBEDL 780  
 Db 721 HNFSPYFTFOSHNSMPGSGREVQIHFDSDQISGLHGNSFHSDEIEYENQKLEEBEDL 780

RESULT 2

US-08-252-517-4  
 ; Sequence 4, Application US/08252517  
 ; Patent No. 5548065  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lemischka, Ihor R.  
 ; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
 ; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
 ; NUMBER OF SEQUENCES: 10  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: ImClone Systems Incorporated  
 ; STREET: 180 Varick Street  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10014  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/252,517  
 ; FILING DATE: 31-OCT-1994  
 ; CLASSIFICATION: 530  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/977,451  
 ; FILING DATE: 19-NOV-1992  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/906,397  
 ; FILING DATE: 26-JUN-1992  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US PCT/US92/05401  
 ; FILING DATE: 26-JUN-1992  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: TW 81102961  
 ; FILING DATE: 15-APR-1992  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US PCT/US92/02750  
 ; FILING DATE: 02-APR-1992  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/813,593  
 ; FILING DATE: 24-DEC-1991  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/793,065  
 ; FILING DATE: 15-NOV-1991  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/728,913  
 ; FILING DATE: 28-JUN-1991  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/679,666  
 ; FILING DATE: 02-APR-1991  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Feit, Irving N.  
 ; REGISTRATION NUMBER: 28,601  
 ; REFERENCE/DOCKET NUMBER: LEM-3-7P

QY 781 NVLTPEDLCCFAYQVAKGMFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDS 840  
 Db 781 NVLTPEDLCCFAYQVAKGMFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDS 840  
 QY 841 NYVVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNPPGIPVDANFYK 900  
 Db 841 NYVVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNPPGIPVDANFYK 900  
 QY 901 LIQNGFKMDOPFYATEEIIYIMQSCWAFDSRKRKPSFNLTSFLGCOLADAEAEAMYQNVDG 960  
 Db 901 LIQNGFKMDOPFYATEEIIYIMQSCWAFDSRKRKPSFNLTSFLGCOLADAEAEAMYQNVDG 960  
 QY 961 RVSECPHTYQNRPPFRSREMDLGLLSPQAQVEDS 993  
 Db 961 RVSECPHTYQNRPPFRSREMDLGLLSPQAQVEDS 993

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 993 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-252-517-4

Query Match 100.0%; Score 5274; DB 1; Length 993;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPALADAGTVPLLVVFSAMIFGTTNODLPVKVLINHKNDSSVGKSSSYPMVSESP 60  
Db 1 MPALADAGTVPLLVVFSAMIFGTTNODLPVKVLINHKNDSSVGKSSSYPMVSESP 60

Qy 61 EDLGCALRQSSGTYEAAAEVVDVSASITLQVLVDAPGNISCLWVFXHSSINCPHFDL 120  
Db 61 EDLGCALRQSSGTYEAAAEVVDVSASITLQVLVDAPGNISCLWVFXHSSINCPHFDL 120

Qy 121 QNRGVSMVILKMTETQAGEYLLFTQSEATNYTLFTVSIIRNTLLYLRPPYFRKMNQD 180  
Db 121 QNRGVSMVILKMTETQAGEYLLFTQSEATNYTLFTVSIIRNTLLYLRPPYFRKMNQD 180

Qy 181 ALVCISESPVPIVEWVLCDSQGSCKEESPAVVKKEKVLHFLGTDIRCARNELGRE 240  
Db 181 ALVCISESPVPIVEWVLCDSQGSCKEESPAVVKKEKVLHFLGTDIRCARNELGRE 240

Qy 241 CTRLTIDNLNTPQTLTLPFLKVGEPWIRKAVVNHGFLTWELNKALEEGNYEM 300  
Db 241 CTRLTIDNLNTPQTLTLPFLKVGEPWIRKAVVNHGFLTWELNKALEEGNYEM 300

Qy 301 STYSTNRMTIRILFAFVSSVARNDGYTCSSSRHPSOSALVTIVGKGFINATNSSEYE 360  
Db 301 STYSTNRMTIRILFAFVSSVARNDGYTCSSSRHPSOSALVTIVGKGFINATNSSEYE 360

Qy 361 IDQYEEFCFSVRKAYPOIRCTWFSRKSFFCEQKGLDNGYSISKFNHKKHQPGYIIPA 420  
Db 361 IDQYEEFCFSVRKAYPOIRCTWFSRKSFFCEQKGLDNGYSISKFNHKKHQPGYIIPA 420

Qy 421 ENDDAQFTKMTLNIIRKQVLAASASQASCFSDGYPLPMTWTKCKSDKSPNCTEITE 480  
Db 421 ENDDAQFTKMTLNIIRKQVLAASASQASCFSDGYPLPMTWTKCKSDKSPNCTEITE 480

Qy 481 GVMNRKANRKFQGVWSSSTLNMSEAIGFLVKCCAYNSLGTSCETILLNSPGPPFIQD 540  
Db 481 GVMNRKANRKFQGVWSSSTLNMSEAIGFLVKCCAYNSLGTSCETILLNSPGPPFIQD 540

Qy 541 NISFYATIGVCLLFTVLTLLICHKYKQFYESQLOMVQVGTSSDNEYFYVDPREYEYD 600  
Db 541 NISFYATIGVCLLFTVLTLLICHKYKQFYESQLOMVQVGTSSDNEYFYVDPREYEYD 600

Qy 601 LKWEFFPRENLFGKVLGSAFGKVMNATAYGISTGVSQVAVKMLKEKADSSEREALMS 660  
Db 601 LKWEFFPRENLFGKVLGSAFGKVMNATAYGISTGVSQVAVKMLKEKADSSEREALMS 660

Qy 661 ELKQMTQLGSHENIVNLLGACTLSGPYILIFEYCCYGDLLNLYRSKREKFRHTWTEIFKE 720  
Db 661 ELKQMTQLGSHENIVNLLGACTLSGPYILIFEYCCYGDLLNLYRSKREKFRHTWTEIFKE 720

Qy 721 HNFSPYPTFQSHPNSSMPGSRVQIHPDSQDQISGLHGNFSHSEDEIEYENOKRLEEDL 780  
Db 721 HNFSPYPTFQSHPNSSMPGSRVQIHPDSQDQISGLHGNFSHSEDEIEYENOKRLEEDL 780

Qy 781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDS 840  
Db 781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDS 840

Qy 841 NYVVRGNARLPVKWMAPESLPEGIYTIKSDVWSYGILLWEIFSLGVNPPGIPVDANFYK 900

Db 841 NYVVRGNARLPVKWMAPESLPEGIYTIKSDVWSYGILLWEIFSLGVNPPGIPVDANFYK 900  
Qy 901 LIQNGFKMDQPFYATERIYIIIMQSCWAFDSRKRSPFNLTSLFGCOLADABEAMYNVDG 960  
Db 901 LIQNGFKMDQPFYATERIYIIIMQSCWAFDSRKRSPFNLTSLFGCOLADABEAMYNVDG 960  
Qy 961 RVSECPHTYQNRPPFSREMDLGLLSPOAQVEDS 993  
Db 961 RVSECPHTYQNRPPFSREMDLGLLSPOAQVEDS 993

RESULT 3  
US-07-906-397A-4  
; Sequence 4, Application US/07906397A  
; Patent No. 5621090  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMCLONE SYSTEMS INCORPORATED  
; STREET: 180 VARICK STREET  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/906.397A  
; FILING DATE: 19920626  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/813,593  
; FILING DATE: 24-DEC-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/793,065  
; FILING DATE: 15-NOV-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/728,913  
; FILING DATE: 28-JUN-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/679,666  
; FILING DATE: 02-APR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Feit, Irving N.  
; REGISTRATION NUMBER: 28,601  
; REFERENCE/DOCKET NUMBER: LEM-3-PPPPPP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-645-1405  
; TELEFAX: 212-645-2054  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 993 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-906-397A-4

Query Match 100.0%; Score 5274; DB 1; Length 993;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPALADAGTVPLLVVFSAMIFGTTNODLPVKVLINHKNDSSVGKSSSYPMVSESP 60  
Db 1 MPALADAGTVPLLVVFSAMIFGTTNODLPVKVLINHKNDSSVGKSSSYPMVSESP 60

Qy 61 EDLGCALRQSSGTYEAAAEVVDVSASITLQVLVDAPGNISCLWVFXHSSINCPHFDL 120

Db 61 EDLGCALPQSSGTVYAAAEVVDVSAITLQVLVDAPGNISCLWFKHSSLNCPHFDL 120  
Qy 121 QNRGVSVMLKMTQAGEYLLFIQSEATNTYTLFTVSRNTLLYTLRRPYFRMENQD 180  
Db 121 QNRGVSVMLKMTQAGEYLLFIQSEATNTYTLFTVSRNTLLYTLRRPYFRMENQD 180  
Qy 181 ALVCISEVPPIVWVLCDSOGESCKEESPAVVKKEKVLHELFGTDIRCCARNEGRE 240  
Db 181 ALVCISEVPPIVWVLCDSOGESCKEESPAVVKKEKVLHELFGTDIRCCARNEGRE 240  
Qy 241 CTRFLTDLNOTPQTLPLKVGCEPLWIRCKAVHVNHGFLTWELNKALEEGNYFEM 300  
Db 241 CTRFLTDLNOTPQTLPLKVGCEPLWIRCKAVHVNHGFLTWELNKALEEGNYFEM 300  
Qy 301 STYSTNRTMIRILFAFVSSVARNDGYTTCSSSKHPSQSALVTIVGKGFINATNSSEDE 360  
Db 301 STYSTNRTMIRILFAFVSSVARNDGYTTCSSSKHPSQSALVTIVGKGFINATNSSEDE 360  
Qy 361 IDQYEEFCFSVRFKAYPOIRCTWTFSRKSFPCEQKGLONGYISIKFCNHKHQPGYIIFHA 420  
Db 361 IDQYEEFCFSVRFKAYPOIRCTWTFSRKSFPCEQKGLONGYISIKFCNHKHQPGYIIFHA 420  
Qy 421 ENDDAQFTKMTLNIRRRPQVLAESAQASCFSDGYPLPSWTWKKKSDKSPNCTEEITE 480  
Db 421 ENDDAQFTKMTLNIRRRPQVLAESAQASCFSDGYPLPSWTWKKKSDKSPNCTEEITE 480  
Qy 481 GVMNRKANRKFVGOWSSSTLNNSBAIKGFLVKCAYNSLGTSCETILLNSPGPPFIQD 540  
Db 481 GVMNRKANRKFVGOWSSSTLNNSBAIKGFLVKCAYNSLGTSCETILLNSPGPPFIQD 540  
Qy 541 NISFYATIGVCLLFIWTLTLCHYKKQFRYESQLQMVQVGTSSDNEYFYVDPREYED 600  
Db 541 NISFYATIGVCLLFIWTLTLCHYKKQFRYESQLQMVQVGTSSDNEYFYVDPREYED 600  
Qy 601 LKWEPPRNLEPGKVLGSGAFKVMNATAYGSKTGVSIQVAVKMLKEKADSSEREALMS 660  
Db 601 LKWEPPRNLEPGKVLGSGAFKVMNATAYGSKTGVSIQVAVKMLKEKADSSEREALMS 660  
Qy 661 ELKMTQLGSHENIYNLLGACTLSGPIYLIFCYCCVGDLLNLRKREKFRHTWTWEIFKE 720  
Db 661 ELKMTQLGSHENIYNLLGACTLSGPIYLIFCYCCVGDLLNLRKREKFRHTWTWEIFKE 720  
Qy 721 HNFSPYPTFQSHPNSSMPGSRVQIHPSDQISGLHNSFHSEDEIYEYENKRLSEEDL 780  
Db 721 HNFSPYPTFQSHPNSSMPGSRVQIHPSDQISGLHNSFHSEDEIYEYENKRLSEEDL 780  
Qy 781 NVLTPEDLLCFAYQVAKGMEFLPKSCVHRDLAARNVLTGKVKVVICDFGLARDIMSDS 840  
Db 781 NVLTPEDLLCFAYQVAKGMEFLPKSCVHRDLAARNVLTGKVKVVICDFGLARDIMSDS 840  
Qy 841 NVVVRGNARLPVKWAPESLFGIYTIKSDVMSYGLLWEIFSLGVNYPGIPVDANFYK 900  
Db 841 NVVVRGNARLPVKWAPESLFGIYTIKSDVMSYGLLWEIFSLGVNYPGIPVDANFYK 900  
Qy 901 LIQNGFKMDQPYATEEIIYIMQSCWAFDSRKRPPNLTSLFGCOLADAEEAMYNQVNDG 960  
Db 901 LIQNGFKMDQPYATEEIIYIMQSCWAFDSRKRPPNLTSLFGCOLADAEEAMYNQVNDG 960  
Qy 961 RVSECPHTYQNRPPSRREMDLGLSPQAQVEDS 993  
Db 961 RVSECPHTYQNRPPSRREMDLGLSPQAQVEDS 993

RESULT 4

US-08-601-891-4  
; Sequence 4, Application US/08601891  
; Patent No. 5747651

; GENERAL INFORMATION:

; APPLICANT: Lemischka, Ithor R.

; TITLE OF INVENTION: TOTIPOENT HEMATOPOIETIC STEM CELL

; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

ADDRESSEE: ImClone Systems Incorporated  
STREET: 180 Varick Street  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10014  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/601,891  
FILING DATE: 15-FEB-1996  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/977,451  
FILING DATE: 19-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/906,397  
FILING DATE: 26-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US PCT/US92/05401  
FILING DATE: 26-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: TW 81102961  
FILING DATE: 15-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US PCT/US92/02750  
FILING DATE: 02-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/813,593  
FILING DATE: 24-DEC-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/793,065  
FILING DATE: 15-NOV-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/728,913  
FILING DATE: 28-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Feit, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 993 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-601-891-4

Query Match 100.0%; Score 5274; DB 1; Length 993;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPALARDAGTVPLLVVFSAMIEGTTITNODLPVKVLINHNKNDSSVGKSSSYPMVSESP 60  
Db 1 MPALARDAGTVPLLVVFSAMIEGTTITNODLPVKVLINHNKNDSSVGKSSSYPMVSESP 60  
Qy 61 EDLGCALPQSSGTVYAAAEVVDVSAITLQVLVDAPGNISCLWFKHSSLNCPHFDL 120  
Db 61 EDLGCALPQSSGTVYAAAEVVDVSAITLQVLVDAPGNISCLWFKHSSLNCPHFDL 120  
Qy 121 QNRGVSVMLKMTQAGEYLLFIQSEATNTYTLFTVSRNTLLYTLRRPYFRMENQD 180  
Db 121 QNRGVSVMLKMTQAGEYLLFIQSEATNTYTLFTVSRNTLLYTLRRPYFRMENQD 180

181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFLGTDIRCCARNELGRE 240  
 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFLGTDIRCCARNELGRE 240  
 241 CTRLFTIDLNOTPTTLQPLFLKVGEPILWIRKAVHVNHGGLTWELNKALEBGNFYFEM 300  
 241 CTRLFTIDLNOTPTTLQPLFLKVGEPILWIRKAVHVNHGGLTWELNKALEBGNFYFEM 300  
 301 STYSTNRMTIRILPAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFINATNSSDEYE 360  
 301 STYSTNRMTIRILPAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFINATNSSDEYE 360  
 361 IDOYEFCFVSFRKAYQIRCTWFSKSPCEOKGLDNGYSISKFNHKKHPOGEYIFHA 420  
 361 IDOYEFCFVSFRKAYQIRCTWFSKSPCEOKGLDNGYSISKFNHKKHPOGEYIFHA 420  
 421 ENDDAQFTKFTLNIRKPKQVLAESAASQASCFSDGYPLPSWTWKCKSDKSPNCTEITE 480  
 421 ENDDAQFTKFTLNIRKPKQVLAESAASQASCFSDGYPLPSWTWKCKSDKSPNCTEITE 480  
 481 GVNVRKANRKFQGVWSSSTLNSEAIKGLVKCAYNLSGTSCTETILLNSPGPPFIQD 540  
 481 GVNVRKANRKFQGVWSSSTLNSEAIKGLVKCAYNLSGTSCTETILLNSPGPPFIQD 540  
 541 NISFYATIGVCLLFIVVLTLLICHKYKQRYESQLOMVQVGTSSDNEYFYVDFREYED 600  
 541 NISFYATIGVCLLFIVVLTLLICHKYKQRYESQLOMVQVGTSSDNEYFYVDFREYED 600  
 601 LWKEFFRENLEFGKVLGSGAGFKVMNATAYGISTGVSIOAVKMLKEKADSSEREALMS 660  
 601 LWKEFFRENLEFGKVLGSGAGFKVMNATAYGISTGVSIOAVKMLKEKADSSEREALMS 660  
 661 ELQWMTQLGSHENIVNLGACTLSGPTLYLFEYCCYGDLLNLYRSKREKTHRTWTEIFKE 720  
 661 ELQWMTQLGSHENIVNLGACTLSGPTLYLFEYCCYGDLLNLYRSKREKTHRTWTEIFKE 720  
 721 HNFSEYPTFOSHNSPPGSGREVOIHDPDSQISGLHGNPSHSEDEIEYENQKLEBEDL 780  
 721 HNFSEYPTFOSHNSPPGSGREVOIHDPDSQISGLHGNPSHSEDEIEYENQKLEBEDL 780  
 781 NVLTFEDLLCFAYQVAKGMEFLBPKSCVHRDLAARNVLVTHGKVKVVICDFGLARDINSDS 840  
 781 NVLTFEDLLCFAYQVAKGMEFLBPKSCVHRDLAARNVLVTHGKVKVVICDFGLARDINSDS 840  
 841 NYVVRGNARLPVKWMAPESLFEGITYTIKSDVWSYGILLWEIFSLGVNPNYPGIPVDANFYK 900  
 841 NYVVRGNARLPVKWMAPESLFEGITYTIKSDVWSYGILLWEIFSLGVNPNYPGIPVDANFYK 900  
 901 LIQNGFMDQPPYATEEIIYIMOSWAFDGRKPSFNLTSLFLGCQLADABEAMYQNVGD 960  
 901 LIQNGFMDQPPYATEEIIYIMOSWAFDGRKPSFNLTSLFLGCQLADABEAMYQNVGD 960  
 961 RVSECPHTYQNRPPFREMDELGLLSPOAQVEDS 993  
 961 RVSECPHTYQNRPPFREMDELGLLSPOAQVEDS 993

RESULT 5  
 US-09-021-324-4  
 ; Sequence 4, Application US/09021324  
 ; Patent No. 5912133  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lemischka, Ibor R.  
 ; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
 ; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
 ; NUMBER OF SEQUENCES: 10  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: ImClone Systems Incorporated  
 ; STREET: 180 Varick Street  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10014

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/021,324  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/07/977,451  
 FILING DATE: 1992-11-19  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/906,397  
 FILING DATE: 26-JUN-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US PCT/US92/05401  
 FILING DATE: 26-JUN-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: TW 81102961  
 FILING DATE: 15-APR-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US PCT/US92/02750  
 FILING DATE: 02-APR-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/813,593  
 FILING DATE: 24-DEC-1991  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/793,065  
 FILING DATE: 15-NOV-1991  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/728,913  
 FILING DATE: 28-JUN-1991  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/679,666  
 FILING DATE: 02-APR-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Felt, Irving N.  
 REGISTRATION NUMBER: 28,601  
 REFERENCE/DOCKET NUMBER: LEM-3-7P  
 TELEPHONE: 212-645-1405  
 TELEFAX: 212-645-2054  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 993 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein

Query Match 100.0%; Score 5274; DB 2; Length 993;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPALARDAGTVPLLVVFSAIFGTITNQDLFVKVLINHNKNDSSVGKSSSYPMVSESP 60  
 Db 1 MPALARDAGTVPLLVVFSAIFGTITNQDLFVKVLINHNKNDSSVGKSSSYPMVSESP 60  
 Qy 61 EDLGCALRPOSSGTVYEAANAEEVDVDSASITLQVLVDAPGNISCLWVPKHSLSNCQPHFDL 120  
 Db 61 EDLGCALRPOSSGTVYEAANAEEVDVDSASITLQVLVDAPGNISCLWVPKHSLSNCQPHFDL 120  
 Qy 121 QNRGVVSWILKMTETQAGEYLLFIQSEATNYTLFTVSRNTLLYTLRBPYFKMENQD 180  
 Db 121 QNRGVVSWILKMTETQAGEYLLFIQSEATNYTLFTVSRNTLLYTLRBPYFKMENQD 180  
 Qy 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFLGTDIRCCARNELGRE 240  
 Db 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFLGTDIRCCARNELGRE 240  
 Qy 241 CTRLFTIDLNOTPTTLQPLFLKVGEPILWIRKAVHVNHGGLTWELNKALEBGNFYFEM 300

```

Db 241 CTLRFTIDLNTQTTLTLPQLFLKVGEPWIRCKAVHVNHGFGLTWELNKALEEGNVFEM 300
Qy 301 STYSTNRTRMIRILFAFVSSVARNDGTGYTCSSSKHPSQSALVTIVGKGFNATNSSDEYE 360
Db 301 STYSTNRTRMIRILFAFVSSVARNDGTGYTCSSSKHPSQSALVTIVGKGFNATNSSDEYE 360
Qy 361 IDQYEEFCFSVRPKAYPQIRCTWTFGRKSPCKQKGLDNGYSISKFCNKHQPGYEYIFHA 420
Db 361 IDQYEEFCFSVRPKAYPQIRCTWTFGRKSPCKQKGLDNGYSISKFCNKHQPGYEYIFHA 420
Qy 421 ENDDAQFTKMTLNRKPKQVLAESAASQSCFSDGYPLPSWTWKCKSDKSPNCTEITE 480
Db 421 ENDDAQFTKMTLNRKPKQVLAESAASQSCFSDGYPLPSWTWKCKSDKSPNCTEITE 480
Qy 481 GWNRRKANRVFGQWSSSTLNSEAIKGPLVKCCAYNSLGTSCETILLNSPGPFPIQD 540
Db 481 GWNRRKANRVFGQWSSSTLNSEAIKGPLVKCCAYNSLGTSCETILLNSPGPFPIQD 540
Qy 541 NISFYATIGVCLLFIVVLTLLICHKYKQPRYSOLOQVGTSSDNEYFYVDFREYED 600
Db 541 NISFYATIGVCLLFIVVLTLLICHKYKQPRYSOLOQVGTSSDNEYFYVDFREYED 600
Qy 601 LWKEPPRENLEFGKVLGSGAFGKVMNATAYGISKTGVSIOVAVKMLKEKADSSERREALMS 660
Db 601 LWKEPPRENLEFGKVLGSGAFGKVMNATAYGISKTGVSIOVAVKMLKEKADSSERREALMS 660
Qy 661 ELKQMTQLGSHENIVNLLGACTLSGPIYLIPEYCCYGDLLNLYRSKREKPHRTWTEIFKE 720
Db 661 ELKQMTQLGSHENIVNLLGACTLSGPIYLIPEYCCYGDLLNLYRSKREKPHRTWTEIFKE 720
Qy 721 HNFSPYTFQSHNSMPPGSRVQIHPDSQISGLHGNFSHSEDEIYENOKRLEEBEDL 780
Db 721 HNFSPYTFQSHNSMPPGSRVQIHPDSQISGLHGNFSHSEDEIYENOKRLEEBEDL 780
Qy 781 NVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIMSDS 840
Db 781 NVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIMSDS 840
Qy 841 NYVVRGNARLPVKWMAPESEFEGIYTIKSDVMSYGILLWEIFSLGVNPNYPGIPVDANFYK 900
Db 841 NYVVRGNARLPVKWMAPESEFEGIYTIKSDVMSYGILLWEIFSLGVNPNYPGIPVDANFYK 900
Qy 901 LIQNGFKMDQPFYATEIYIIMOSCAFDGRKPSFNLTSFLGCQLADAEAMQYVNDG 960
Db 901 LIQNGFKMDQPFYATEIYIIMOSCAFDGRKPSFNLTSFLGCQLADAEAMQYVNDG 960
Qy 961 RVSECPHTYQNRPPFSREMDLGLLSPQAQVEDS 993
Db 961 RVSECPHTYQNRPPFSREMDLGLLSPQAQVEDS 993

```

RESULT 6

```

PCT-US92-09893-4
; Sequence 4, Application PC/TUS9209893
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/09893

```

```

; FILING DATE: 19921116
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Felt, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-7PT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 993 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US92-09893-4

Query Match 100.0%; Score 5274; DB 5; Length 993;
Best Local Similarity 100.0%; Pred No. 0;
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPALARDAGTVPLLVVFSAIFGTITNQDLPVTKCVLINHKNDSSVKGKSSYPMVSESP 60
Db 1 MPALARDAGTVPLLVVFSAIFGTITNQDLPVTKCVLINHKNDSSVKGKSSYPMVSESP 60
Qy 61 EDLGCALRPOSSGTVEAAAEVDVSASITLQVLVDAPGNISCLWVFKHSLNCQPHFDL 120
Db 61 EDLGCALRPOSSGTVEAAAEVDVSASITLQVLVDAPGNISCLWVFKHSLNCQPHFDL 120
Qy 121 QNGGVVSWMLTKMTQAGYLLFIOSEATNYTILFTVSRINTLLYTLRRPYFRMENQD 180
Db 121 QNGGVVSWMLTKMTQAGYLLFIOSEATNYTILFTVSRINTLLYTLRRPYFRMENQD 180
Qy 181 ALVCISESPPEPVEVWVLCDSQGESCKEESPAVVKKEKVLHFLGTDIRCCARNELGRE 240
Db 181 ALVCISESPPEPVEVWVLCDSQGESCKEESPAVVKKEKVLHFLGTDIRCCARNELGRE 240
Qy 241 CTLRFTIDLNTQTTLTLPQLFLKVGEPWIRCKAVHVNHGFGLTWELNKALEEGNVFEM 300
Db 241 CTLRFTIDLNTQTTLTLPQLFLKVGEPWIRCKAVHVNHGFGLTWELNKALEEGNVFEM 300
Qy 301 STYSTNRTRMIRILFAFVSSVARNDGTGYTCSSSKHPSQSALVTIVGKGFNATNSSDEYE 360
Db 301 STYSTNRTRMIRILFAFVSSVARNDGTGYTCSSSKHPSQSALVTIVGKGFNATNSSDEYE 360
Qy 361 IDQYEEFCFSVRPKAYPQIRCTWTFGRKSPCKQKGLDNGYSISKFCNKHQPGYEYIFHA 420
Db 361 IDQYEEFCFSVRPKAYPQIRCTWTFGRKSPCKQKGLDNGYSISKFCNKHQPGYEYIFHA 420
Qy 421 ENDDAQFTKMTLNRKPKQVLAESAASQSCFSDGYPLPSWTWKCKSDKSPNCTEITE 480
Db 421 ENDDAQFTKMTLNRKPKQVLAESAASQSCFSDGYPLPSWTWKCKSDKSPNCTEITE 480
Qy 481 GWNRRKANRVFGQWSSSTLNSEAIKGPLVKCCAYNSLGTSCETILLNSPGPFPIQD 540
Db 481 GWNRRKANRVFGQWSSSTLNSEAIKGPLVKCCAYNSLGTSCETILLNSPGPFPIQD 540
Qy 541 NISFYATIGVCLLFIVVLTLLICHKYKQPRYSOLOQVGTSSDNEYFYVDFREYED 600
Db 541 NISFYATIGVCLLFIVVLTLLICHKYKQPRYSOLOQVGTSSDNEYFYVDFREYED 600
Qy 601 LWKEPPRENLEFGKVLGSGAFGKVMNATAYGISKTGVSIOVAVKMLKEKADSSERREALMS 660
Db 601 LWKEPPRENLEFGKVLGSGAFGKVMNATAYGISKTGVSIOVAVKMLKEKADSSERREALMS 660
Qy 661 ELKQMTQLGSHENIVNLLGACTLSGPIYLIPEYCCYGDLLNLYRSKREKPHRTWTEIFKE 720
Db 661 ELKQMTQLGSHENIVNLLGACTLSGPIYLIPEYCCYGDLLNLYRSKREKPHRTWTEIFKE 720
Qy 721 HNFSPYTFQSHNSMPPGSRVQIHPDSQISGLHGNFSHSEDEIYENOKRLEEBEDL 780
Db 721 HNFSPYTFQSHNSMPPGSRVQIHPDSQISGLHGNFSHSEDEIYENOKRLEEBEDL 780

```

781 NVLTFFDLCCFAYQVAKGMEFLKSCVHRDLAARNVLVTHGKVKICDFGLARDIMSDS 840  
 781 NVLTFFDLCCFAYQVAKGMEFLKSCVHRDLAARNVLVTHGKVKICDFGLARDIMSDS 840  
 841 NYVVRGNARLPVKWMAPELSFEGYITIKSDVMSYGILLWEIFSLGVNYPGIPVDANFYK 900  
 841 NYVVRGNARLPVKWMAPELSFEGYITIKSDVMSYGILLWEIFSLGVNYPGIPVDANFYK 900  
 901 LIQNGFKMDQPPYATEEIIYIMQSCWAFDSKRSPFNLTSLFLGCOLADAEAMYQNVGD 960  
 901 LIQNGFKMDQPPYATEEIIYIMQSCWAFDSKRSPFNLTSLFLGCOLADAEAMYQNVGD 960  
 961 RVSECPHTYQNRPRPFSREMDLGLLSPQAQVEDS 993  
 961 RVSECPHTYQNRPRPFSREMDLGLLSPQAQVEDS 993

RESULT 7  
 PCT-US92-05401-4  
 ; Sequence 4, Application PC/TUS9205401  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lemischka, Ihor R.  
 ; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
 ; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: INCLONE SYSTEMS INCORPORATED  
 ; STREET: 180 VARICK STREET  
 ; CITY: NEW YORK  
 ; STATE: NEW YORK  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10014  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US92/05401  
 ; FILING DATE: 19920626  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Feit, Irving N.  
 ; REGISTRATION NUMBER: 28,601  
 ; REFERENCE/DOCKET NUMBER: LEM-3-PPPPPT  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 212-645-1405  
 ; TELEFAX: 212-645-2054  
 ; INFORMATION FOR SEQ ID NO. 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1160 amino acids  
 ; TYPE: AMINO ACID  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; PCT-US92-05401-4

Query Match 100.0%; Score 5274; DB 5; Length 1160;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MPALARDAGTVPLLVFVSAMIFGTITNQDLFVKVCLINHNKNDSSVGVKSSSYPMWSESP 60  
 DB 20 MPALARDAGTVPLLVFVSAMIFGTITNQDLFVKVCLINHNKNDSSVGVKSSSYPMWSESP 79  
 QY 61 EDGCLARPOSSGTVVEAAVEVDVSATITQVLVDAPGNISCLWVFKHSLNCQPHFDL 120  
 DB 80 EDGCLARPOSSGTVVEAAVEVDVSATITQVLVDAPGNISCLWVFKHSLNCQPHFDL 139  
 QY 121 QNRGVVSWILKMTQAGYLLFIQSEATNYTLFTVSVIRNTLLYTLRRPYFRKMNQD 180  
 DB 140 QNRGVVSWILKMTQAGYLLFIQSEATNYTLFTVSVIRNTLLYTLRRPYFRKMNQD 199  
 QY 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240

200 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 259  
 241 CTRLFTIDLNTPTQLFLKVGELWIRCKA VVNVHFGGLTWELNKALEEGNYPFM 300  
 260 CTRLFTIDLNTPTQLFLKVGELWIRCKA VVNVHFGGLTWELNKALEEGNYPFM 319  
 301 STYSTNRTMIRILFAFVSSVARNDTYTSCSSKHPQSALVTIIVKGFINATNSSEDEYE 360  
 320 STYSTNRTMIRILFAFVSSVARNDTYTSCSSKHPQSALVTIIVKGFINATNSSEDEYE 379  
 361 IDQYEEFCFSVRFKAYPOIRCTWTFPSRKSFPCEQKGLDNGYSISKFCNKHQPGGEYIFHA 420  
 380 IDQYEEFCFSVRFKAYPOIRCTWTFPSRKSFPCEQKGLDNGYSISKFCNKHQPGGEYIFHA 439  
 421 ENDDAQFTKMTFLNIRKPKQVLAESAASQSCFSDGYPLPSWTWKCKSDKSPNCTEEITE 480  
 440 ENDDAQFTKMTFLNIRKPKQVLAESAASQSCFSDGYPLPSWTWKCKSDKSPNCTEEITE 499  
 481 GWNRKANRKFVGGWVSSSTILNMSAIKGLVKCCAYNSLGTSCETILLNSPGPPFIQD 540  
 500 GWNRKANRKFVGGWVSSSTILNMSAIKGLVKCCAYNSLGTSCETILLNSPGPPFIQD 559  
 541 NISFYATIGVCLLFTIVLTLIICHYKQKQRYESQLOMVQVGTGSSDNEFYVDFREYED 600  
 560 NISFYATIGVCLLFTIVLTLIICHYKQKQRYESQLOMVQVGTGSSDNEFYVDFREYED 619  
 601 LKWEFPRENLEFGKVLGSGAFKVMNATAYGSKTGVSIQVAVKMLKADSSEREALMS 660  
 620 LKWEFPRENLEFGKVLGSGAFKVMNATAYGSKTGVSIQVAVKMLKADSSEREALMS 679  
 661 ELKMTQLGSHENIVNLLGACTLSGPIYLI FEYCCYGDLLNLYLSKREKFRHTWTETPK 720  
 680 ELKMTQLGSHENIVNLLGACTLSGPIYLI FEYCCYGDLLNLYLSKREKFRHTWTETPK 739  
 721 HNFSFYPTFQSHPNSSMPGSEVQIHDPDSDOISGLHGNSFSEDEIEYENQKREEDDL 780  
 740 HNFSFYPTFQSHPNSSMPGSEVQIHDPDSDOISGLHGNSFSEDEIEYENQKREEDDL 799  
 781 NVLTFFDLCCFAYQVAKGMEFLKSCVHRDLAARNVLVTHGKVKICDFGLARDIMSDS 840  
 800 NVLTFFDLCCFAYQVAKGMEFLKSCVHRDLAARNVLVTHGKVKICDFGLARDIMSDS 859  
 841 NYVVRGNARLPVKWMAPELSFEGYITIKSDVMSYGILLWEIFSLGVNYPGIPVDANFYK 900  
 860 NYVVRGNARLPVKWMAPELSFEGYITIKSDVMSYGILLWEIFSLGVNYPGIPVDANFYK 919  
 901 LIQNGFKMDQPPYATEEIIYIMQSCWAFDSKRSPFNLTSLFLGCOLADAEAMYQNVGD 960  
 920 LIQNGFKMDQPPYATEEIIYIMQSCWAFDSKRSPFNLTSLFLGCOLADAEAMYQNVGD 979  
 961 RVSECPHTYQNRPRPFSREMDLGLLSPQAQVEDS 993  
 980 RVSECPHTYQNRPRPFSREMDLGLLSPQAQVEDS 1012

RESULT 8  
 US-08-222-299-4  
 ; Sequence 4, Application US/08222299  
 ; Patent No. 5635388  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bennett, Brian D.  
 ; APPLICANT: Broz, Susan D.  
 ; APPLICANT: Matthews, William  
 ; APPLICANT: Zeigler, Francis C.  
 ; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES TH  
 ; NUMBER OF SEQUENCES: 4  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Genentech, Inc.  
 ; STREET: 460 Point San Bruno Blvd  
 ; CITY: South San Francisco  
 ; STATE: California  
 ; COUNTRY: USA

Query Match 99.4%; Score 5240; DB 1; Length 993;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 98; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPILVVSAMIFGTITNQDLPIKCVLINHKNDSSVSKSSSYPMVSESP 60  
DB 1 MPALARDGGQLPLLVVSAMIFGTITNQDLPIKCVLINHKNDSSVSKSSSYPMVSESP 60

QY 61 EDLGCALRPOSSGTYVAAAUEVDVASITLQVLVDAPGNI SCLWVFKHSLNCOQPHDL 120  
DB 61 EDLGCALRPOSSGTYVAAAUEVDVASITLQVLVDAPGNI SCLWVFKHSLNCOQPHDL 120

QY 121 QNKGVSWMILKMTETQAGEVLLFIQSEATNTTLFTVSRINTLLYTLRRPYRKMENQD 180  
DB 121 QNKGVSWMILKMTETQAGEVLLFIQSEATNTTLFTVSRINTLLYTLRRPYRKMENQD 180

QY 181 ALVCISESVPEPIVWVLCDSQGSCKEESPAVVKKEKVLHFGTDIRCCARNELGRE 240  
DB 181 ALVCISESVPEPIVWVLCDSQGSCKEESPAVVKKEKVLHFGMDIRCCARNELGRE 240

QY 241 CTRLFTIDLNOTPTTLPOLFLKVGEPILWIRCKAVHNVHGGFLTWELNKALESNGYFEM 300  
DB 241 CTRLFTIDLNOTPTTLPOLFLKVGEPILWIRCKAVHNVHGGFLTWELNKALESNGYFEM 300

QY 301 STYSTNRTMIRILFAFVSSVARNDGTGYTCSSKHPSOSALVTIVGKGFINATNSSEDE 360  
DB 301 STYSTNRTMIRILFAFVSSVARNDGTGYTCSSKHPSOSALVTIVGKGFINATNSSEDE 360

QY 361 IDQVEEFCFSVRKAYPOIRCTWTFRSKSPCEQKGLDNGYSISKFCNKHQPOEYIFHA 420  
DB 361 IDQVEEFCFSVRKAYPOIRCTWTFRSKSPCEQKGLDNGYSISKFCNKHQPOEYIFHA 420

QY 421 ENDDAQTKMFTLNIRKQVLABASASQASCFSDGYPLPSWTWKCKSDKSPNCTEITE 480  
DB 421 ENDDAQTKMFTLNIRKQVLABASASQASCFSDGYPLPSWTWKCKSDKSPNCTEITE 480

QY 481 GWNRKANRKFVGOVSSSTLNWSEAIKGLVKCCAYNSLGTSCETILLNSPGPPFIQD 540  
DB 481 GWNRKANRKFVGOVSSSTLNWSEAIKGLVKCCAYNSLGTSCETILLNSPGPPFIQD 540

QY 541 NISFYATIGVCLLFIIVTLILCHYKKQFYESQLOQVQVTSDDNEYFYVDPREVEYD 600  
DB 541 NISFYATIGVCLLFIIVTLILCHYKKQFYESQLOQVQVTSDDNEYFYVDPREVEYD 600

QY 601 LKWEFFRENLEFGKVLGSGAFGKVMNATAYGISKTGVSIOQAVKMLKEKADSSREALMS 660

DB 601 LKWEFFRENLEFGKVLGSGAFGKVMNATAYGISKTGVSIOQAVKMLKEKADSSREALMS 660  
QY 661 ELKQMTOLGSHENIVNLLGACTLSGPIYLIFFEYCCYGDLLNLYLSRKEKHFRTWTEIFKE 720  
DB 661 ELKQMTOLGSHENIVNLLGACTLSGPIYLIFFEYCCYGDLLNLYLSRKEKHFRTWTEIFKE 720  
QY 721 HNFSPYPTQSHNSNMPGSRREVQIHPDSDOISGLHGNSEHSEDEIEYENOKRLEEBEDL 780  
DB 721 HNFSPYPTQSHNSNMPGSRREVQIHPDSDOISGLHGNSEHSEDEIEYENOKRLEEBEDL 780  
QY 781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDINSDS 840  
DB 781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDINSDS 840  
QY 841 NYVVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWPIPSLGVPNPYGPVPDANPYK 900  
DB 841 NYVVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWPIPSLGVPNPYGPVPDANPYK 900  
QY 901 LIQNGFKMDOPFYATEBIIYIMQSWAFDSKRPSPFNLTSLFLGCOLADAEAMYNQVNDG 960  
DB 901 LIQNGFKMDOPFYATEBIIYIMQSWAFDSKRPSPFNLTSLFLGCOLADAEAMYNQVNDG 960  
QY 961 RVSECPHTYONRRPFSREMDLGLLSPOAQVEDS 993  
DB 961 PVSECPHTYONRRPFSREMDLGLLSPOAQVEDS 993

RESULT 9  
US-08-434-878-4  
; Sequence 4, Application US/08434878  
; Patent No. 5997865  
; GENERAL INFORMATION:  
; APPLICANT: Bennett, Brian D.  
; APPLICANT: Broz, Susan D.  
; APPLICANT: Matthews, William  
; APPLICANT: Zeigler, Francis C.  
; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES TH  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/434,878  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hasak, Janet E.  
; REGISTRATION NUMBER: 28,616  
; REFERENCE/DOCKET NUMBER: 879  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-1896  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 993 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-434-878-4



Query Match 99.4%; Score 5240; DB 2; Length 993;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 987; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MPALARDAGTVPPLLVFSAIFGTITNQDLFVKVLINHNKNDSSVGKSSSYPMVSESP 60  
Db 1 MPALARDGGQLPLLVFSAIFGTITNQDLFVKVLINHNKNDSSVGKSSSYPMVSESP 60

Qy 61 EDLGCALRPOSSGTVYEAATAVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCOPHFDL 120  
Db 61 EDLGCALRPOSSGTVYEAATAVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCOPHFDL 120

Qy 121 QNRGWSMWILKMTETQAGEYLLFIQSEATNYTILFTVSRINTLLYTLRRPYFRMENQD 180  
Db 121 QNRGWSMWILKMTETQAGEYLLFIQSEATNYTILFTVSRINTLLYTLRRPYFRMENQD 180

Qy 181 ALVCISSEVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHELFGMDIRCCARNELGRE 240  
Db 181 ALVCISSEVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHELFGMDIRCCARNELGRE 240

Qy 241 CTRFLTIDLNOTPOTTLFQVGEPLWIRCKAVHNVHGFGLTWELNKALEEGNYFEM 300  
Db 241 CTRFLTIDLNOTPOTTLFQVGEPLWIRCKAVHNVHGFGLTWELNKALEEGNYFEM 300

Qy 301 STYSTNRTRMIRILFAPVSSVARNDTGYTTCSSSKHPSQSALVTIVGKGFINATNSSDEYE 360  
Db 301 STYSTNRTRMIRILFAPVSSVARNDTGYTTCSSSKHPSQSALVTIVGKGFINATNSSDEYE 360

Qy 361 IDOYEEFCFSVRKAYPOIRCTWTFRSKSPCEOKGLDNGYSISKFCNHHKQPGYEYIFHA 420  
Db 361 IDOYEEFCFSVRKAYPOIRCTWTFRSKSPCEOKGLDNGYSISKFCNHHKQPGYEYIFHA 420

Qy 421 ENDDAQTFTKFTNIRKPKQVLAASASQSCFSDGYPLPSWTWKKSDKSPNCTEITE 480  
Db 421 ENDDAQTFTKFTNIRKPKQVLAASASQSCFSDGYPLPSWTWKKSDKSPNCTEITE 480

Qy 481 GVNRRKANRVFGOWSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPFPIQD 540  
Db 481 GVNRRKANRVFGOWSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPFPIQD 540

Qy 541 NISFYATIGVCLLFIVVLTLLIHKYKQRYESQLOQVQVQVQSSDNEYFYVDPREYED 600  
Db 541 NISFYATIGVCLLFIVVLTLLIHKYKQRYESQLOQVQVQVQSSDNEYFYVDPREYED 600

Qy 601 LKWEFFPRENLEFGKVLGSGAFGKVMNATAYGISTGVSIOVAVVOMLKEKADSSREALMS 660  
Db 601 LKWEFFPRENLEFGKVLGSGAFGKVMNATAYGISTGVSIOVAVVOMLKEKADSSREALMS 660

Qy 661 ELKQMTOLGSHENIVNLGACTLSGPIYLIPEYCCYGDLLNLYRSKREKGRHTWTEIFKE 720  
Db 661 ELKQMTOLGSHENIVNLGACTLSGPIYLIPEYCCYGDLLNLYRSKREKGRHTWTEIFKE 720

Qy 721 HNF5FYPTFQSHPNSSMPGSRVQIHDPDSQOISGLHGNFSHSEDEIEYENQKLEEEEDL 780  
Db 721 HNF5FYPTFQSHPNSSMPGSRVQIHDPDSQOISGLHGNFSHSEDEIEYENQKLEEEEDL 780

Qy 781 NVLTFEDLLCFAYQVAKGMFLEPKSCVHRDLAARNVLTGHKVVKICDFGLARDINSDS 840  
Db 781 NVLTFEDLLCFAYQVAKGMFLEPKSCVHRDLAARNVLTGHKVVKICDFGLARDINSDS 840

Qy 841 NYVVRGNARLPVKWMAPELIFEGYITTKSDVMSYGILLWEIFSLGVNYPGIPVDANFYK 900  
Db 841 NYVVRGNARLPVKWMAPELIFEGYITTKSDVMSYGILLWEIFSLGVNYPGIPVDANFYK 900

Qy 901 LIQNGFMDOPFYATEIYIIMOSCAFDGRKSPFNLTSLFCQLADAEAEAMQVNDG 960  
Db 901 LIQNGFMDOPFYATEIYIIMOSCAFDGRKSPFNLTSLFCQLADAEAEAMQVNDG 960

Qy 961 RVSECPHTYQNRPPFSREMDLGLLSPQAQVEDS 993  
Db 961 RVSECPHTYQNRPPFSREMDLGLLSPQAQVEDS 993

RESULT 10  
PCT-US95-03718-4  
; Sequence 4, Application PC/TUS9503718  
; GENERAL INFORMATION:  
; APPLICANT: GENENTECH, INC.  
; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES T  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/03718  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Wendy M. Lee  
; REGISTRATION NUMBER: 00,000  
; REFERENCE/DOCKET NUMBER: 879PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-1994  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 993 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
PCT-US95-03718-4

Query Match 99.4%; Score 5240; DB 5; Length 993;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 987; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MPALARDAGTVPPLLVFSAIFGTITNQDLFVKVLINHNKNDSSVGKSSSYPMVSESP 60  
Db 1 MPALARDGGQLPLLVFSAIFGTITNQDLFVKVLINHNKNDSSVGKSSSYPMVSESP 60

Qy 61 EDLGCALRPOSSGTVYEAATAVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCOPHFDL 120  
Db 61 EDLGCALRPOSSGTVYEAATAVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCOPHFDL 120

Qy 121 QNRGWSMWILKMTETQAGEYLLFIQSEATNYTILFTVSRINTLLYTLRRPYFRMENQD 180  
Db 121 QNRGWSMWILKMTETQAGEYLLFIQSEATNYTILFTVSRINTLLYTLRRPYFRMENQD 180

Qy 181 ALVCISSEVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHELFGMDIRCCARNELGRE 240  
Db 181 ALVCISSEVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHELFGMDIRCCARNELGRE 240

Qy 241 CTRFLTIDLNOTPOTTLFQVGEPLWIRCKAVHNVHGFGLTWELNKALEEGNYFEM 300  
Db 241 CTRFLTIDLNOTPOTTLFQVGEPLWIRCKAVHNVHGFGLTWELNKALEEGNYFEM 300

Qy 301 STYSTNRTRMIRILFAPVSSVARNDTGYTTCSSSKHPSQSALVTIVGKGFINATNSSDEYE 360  
Db 301 STYSTNRTRMIRILFAPVSSVARNDTGYTTCSSSKHPSQSALVTIVGKGFINATNSSDEYE 360

Qy 361 IDOYEEFCFSVRKAYPOIRCTWTFRSKSPCEOKGLDNGYSISKFCNHHKQPGYEYIFHA 420  
Db 361 IDOYEEFCFSVRKAYPOIRCTWTFRSKSPCEOKGLDNGYSISKFCNHHKQPGYEYIFHA 420

QY 421 ENDDAQTQKMTLNIRKPKQVLAASASQSCFSDGYPLPSWTWKCKSDKSPNCTEITE 480  
DB 421 ENDDAQTQKMTLNIRKPKQVLAASASQSCFSDGYPLPSWTWKCKSDKSPNCTEITE 480  
QY 481 GVNRRKANRVFGQWSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPPFIQD 540  
DB 481 GVNRRKANRVFGQWSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPPFIQD 540  
QY 541 NISFYATIGVCLLFIVVLTLLI CHYKQKQFYESQOLQWQVGTSSDNEYFYVDREYED 600  
DB 541 NISFYATIGVCLLFIVVLTLLI CHYKQKQFYESQOLQWQVGTSSDNEYFYVDREYED 600  
QY 601 LKWEFFPRENLEFGKVLGSGAFKGMNATAYGISKTGVSIOVAVKMLKEKADSSREALMS 660  
DB 601 LKWEFFPRENLEFGKVLGSGAFKGMNATAYGISKTGVSIOVAVKMLKEKADSSREALMS 660  
QY 661 ELKQMTQLGSHENIVNLLGACTLSGPIYLI FEYCCYGDLLNLYLSKREKPHRTWTEIFKE 720  
DB 661 ELKQMTQLGSHENIVNLLGACTLSGPIYLI FEYCCYGDLLNLYLSKREKPHRTWTEIFKE 720  
QY 721 HNFSFYPTFQSHPNSSMPGSRVQIHDPDSQOISGLHGNSFHSDEIEYENOKRLEEBEDL 780  
DB 721 HNFSFYPTFQSHPNSSMPGSRVQIHDPDSQOISGLHGNSFHSDEIEYENOKRLEEBEDL 780  
QY 781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGKVKVVICDFGLARDIMSDS 840  
DB 781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGKVKVVICDFGLARDIMSDS 840  
QY 841 NYVVRGNARLPVKWMAPESLFEGYITIKSDVMSYGILLWEIFSLGVNYPGIPVDANFYK 900  
DB 841 NYVVRGNARLPVKWMAPESLFEGYITIKSDVMSYGILLWEIFSLGVNYPGIPVDANFYK 900  
QY 901 LIQNGFKMDQFPYATEIYIIIMQSWAFDRKRSPFNLTSLFLGCOLADAEAMYQNVVDG 960  
DB 901 LIQNGFKMDQFPYATEIYIIIMQSWAFDRKRSPFNLTSLFLGCOLADAEAMYQNVVDG 960  
QY 961 RVSECPHTYQNRPFPSRMDLGLLSPQAQVEDS 993  
DB 961 RVSECPHTYQNRPFPSRMDLGLLSPQAQVEDS 993

RESULT 11

US-08-183-211-2  
; Sequence 2, Application US/08183211  
; Patent No. 5618709  
; GENERAL INFORMATION:  
; APPLICANT: Alan M. Gewirtz, Donald Small, Curt I. Civin.  
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEO- TIDES  
; TITLE OF INVENTION: SPECIFIC FOR STR-1 AND METHOD FOR  
; TITLE OF INVENTION: INHIBITING EXPRESSION OF THE STR-1 PROTEIN  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEIDEL GONDA LAVORGNA & MONACO  
; STREET: Suite 1800, Penn Center Plaza  
; City: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: U.S.A.  
; ZIP: 19102  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/183,211  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Monaco, Daniel A.  
; REGISTRATION NUMBER: 30,480

; REFERENCE/DOCKET NUMBER: 3957-15  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-8383  
; TELEFAX: (215) 568-5549  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 993 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; US-08-183-211-2

Query Match 99.3%; Score 5238; DB 1; Length 993;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 991; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 1 MPALARDAGTVPLLVFVSAMIFGTITNODLPVICKVLINHKNDSSVGKSSSYPMVSSSP 60  
DB 1 MPALARDAGTVPLLVFVSAMIFGTITNODLPVICKVLINHKNDSSVGKSSSYPMVSSSP 60  
QY 61 EDLGCALRPOSSGTYEAAAVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCOPHFDL 120  
DB 61 EDLGCALRPOSSGTYEAAAVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCOPHFDL 120  
QY 121 QNRGVSMVILKMTETOAGEYLLFIQSEATNYTILFTVISIRNTLLYTLRRPYFRKXENOD 180  
DB 121 QNRGVSMVILKMTETOAGEYLLFIQSEATNYTILFTVISIRNTLLYTLRRPYFRKXENOD 180  
QY 181 ALVCISSEVPEPIVEWVLCDSSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240  
DB 181 ALVCISSEVPEPIVEWVLCDSSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240  
QY 241 CTRLTFTDLNQTPTTLPQLFKVGEPLWIRCKAVVNHGFLTWELNKALEEGNYFEM 300  
DB 241 CTRLTFTDLNQTPTTLPQLFKVGEPLWIRCKAVVNHGFLTWELNKALEEGNYFEM 300  
QY 301 STYSTNRTMIRILFAFVSSVARNDTGYTTCSSSHKPSQSALVTIVKGFINATNSSEDEYE 360  
DB 301 STYSTNRTMIRILFAFVSSVARNDTGYTTCSSSHKPSQSALVTIVKGFINATNSSEDEYE 360  
QY 361 IDOYEEFCFSVRFKAYPOIRCTWTFSRKSPFCEQKGLDNGYSISKFCNHKQHPGEYIFH- 419  
DB 361 IDOYEEFCFSVRFKAYPOIRCTWTFSRKSPFCEQKGLDNGYSISKFCNHKQHPGEYIFH- 420  
QY 420 AENDDAQTQKMTLNIRKPKQVLAASASQSCFSDGYPLPSWTWKCKSDKSPNCTEITE 479  
DB 420 AENDDAQTQKMTLNIRKPKQVLAASASQSCFSDGYPLPSWTWKCKSDKSPNCTEITE 479  
QY 480 EGVNRRKANRVFGQWSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPPFIQ 539  
DB 480 EGVNRRKANRVFGQWSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPPFIQ 539  
QY 540 DNISFYATIGVCLLFIVVLTLLI CHYKQKQFYESQOLQWQVGTSSDNEYFYVDREYED 599  
DB 540 DNISFYATIGVCLLFIVVLTLLI CHYKQKQFYESQOLQWQVGTSSDNEYFYVDREYED 599  
QY 600 DLKWEFFPRENLEFGKVLGSGAFKGMNATAYGISKTGVSIOVAVKMLKEKADSSREALM 659  
DB 600 DLKWEFFPRENLEFGKVLGSGAFKGMNATAYGISKTGVSIOVAVKMLKEKADSSREALM 659  
QY 660 SELKQMTQLGSHENIVNLLGACTLSGPIYLI FEYCCYGDLLNLYLSKREKPHRTWTEIFK 719  
DB 660 SELKQMTQLGSHENIVNLLGACTLSGPIYLI FEYCCYGDLLNLYLSKREKPHRTWTEIFK 719  
QY 720 EHNFSFYPTFQSHPNSSMPGSRVQIHDPDSQOISGLHGNSFHSDEIEYENOKRLEEBED 779  
DB 720 EHNFSFYPTFQSHPNSSMPGSRVQIHDPDSQOISGLHGNSFHSDEIEYENOKRLEEBED 779  
QY 780 LNVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGKVKVVICDFGLARDIMSD 839  
DB 780 LNVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGKVKVVICDFGLARDIMSD 839  
QY 840 SNYVVRGNARLPVKWMAPESLFEGYITIKSDVMSYGILLWEIFSLGVNYPGIPVDANFY 899

```

Db      840 SNYVVRGNARLPVKWAPESLFEIGYTIKSDVMSYGILLWEIFSLGVNPPYGPVDANFY 899
QY      900 KLIQNGFMDQPPFATEEIIYIMQSCWAFDSKRKPSFNLTSFLGCOLADAEAMQVND 959
Db      900 KLIQNGFMDQPPFATEEIIYIMQSCWAFDSKRKPSFNLTSFLGCOLADAEAMQVND 959
QY      960 GRVSECPHTYQNRPPFSREMDLGLLSPQAQVEDS 993
Db      960 GRVSECPHTYQNRPPFSREMDLGLLSPQAQVEDS 993

RESULT 12
PCT-US95-00176A-2
; Sequence 2, Application PC/TUS9500176A
; GENERAL INFORMATION:
; APPLICANT: Alan M. Gewirtz, Donald Small, Curt I. Civin.
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEIDEL GONDA LAVORGNA & MONACO
; STREET: Suite 1800, Penn Center Plaza
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 KB
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/00176A
; FILING DATE: 6 January 1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/183,211
; FILING DATE: 14 January 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 3957-14 PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: None
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 993 amino acids
; TYPE: amino acid
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
; PCT-US95-00176A-2

Query Match          99.3%; Score 5238; DB 5; Length 993;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 991; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY      1 MPALARDAGTVPPLLVFVSAMIFGTITNODLPVTKCVLINHKNDSSVGKSSSYPMVSESP 60
Db      1 MPALARDAGTVPPLLVFVSAMIFGTITNODLPVTKCVLINHKNDSSVGKSSSYPMVSESP 60
QY      61 EDLGCALRPQSSGTVEAAAVEVDVSASITLQVLVDAPGNISCLWVFKSSLNCQPHFDL 120
Db      61 EDLGCALRPQSSGTVEAAAVEVDVSASITLQVLVDAPGNISCLWVFKSSLNCQPHFDL 120
QY      121 QNRGVVSMILKMTQAGEYLLFIQSEATNYTLFTVTSIRNTLLYTLRPPYFRKMNQD 180
Db      121 QNRGVVSMILKMTQAGEYLLFIQSEATNYTLFTVTSIRNTLLYTLRPPYFRKMNQD 180
QY      181 ALVCISESPPEIVWVLCDSQGESKESPAVVKKEKVLHELFGTDIRCCARNELGRE 240

```

```

Db      181 ALVCISESPPEIVWVLCDSQGESKESPAVVKKEKVLHELFGTDIRCCARNELGRE 240
QY      241 CTRLTFTDLNQTPTTLTLPOLFLKVGEPWLIRKAVVHVHGFGLTWELFNKALEEGNYFPM 300
Db      241 CTRLTFTDLNQTPTTLTLPOLFLKVGEPWLIRKAVVHVHGFGLTWELFNKALEEGNYFPM 300
QY      301 STYSTNRMTIRILFAFVSSVARNDTYTCCSSSKHPQSALVTIVKGFIPINATNSEDYE 360
Db      301 STYSTNRMTIRILFAFVSSVARNDTYTCCSSSKHPQSALVTIVKGFIPINATNSEDYE 360
QY      361 IDQYEEFCFSVRFKAYPQIRCTWTFSRKSPFCEQKGLDNGYSISKFCNHHKHPGEYIFH- 419
Db      361 IDQYEEFCFSVRFKAYPQIRCTWTFSRKSPFCEQKGLDNGYSISKFCNHHKHPGEYIFH- 420
QY      420 AENDDAQFTKFTLNIRRKPOVLAEASASQSCFSDGYPLPSWTWKKCDKSPNCTEET 479
Db      421 AENDDAQFTKFTLNIRRKPOVLAEASASQSCFSDGYPL-SWTWKKCDKSPNCTEET 479
QY      480 EGVNRRKANRVKFGOWSSSTLANMSEAIKGPLVKCCAYNSLGTSCETILLNSPGPPFTQ 539
Db      480 EGVNRRKANRVKFGOWSSSTLANMSEAIKGPLVKCCAYNSLGTSCETILLNSPGPPFTQ 539
QY      540 DNISFYATIGVCLLFIVVLTLLIICHKVKQPRYESQLOMVOVTGSSDNEYFVVDPREY 599
Db      540 DNISFYATIGVCLLFIVVLTLLIICHKVKQPRYESQLOMVOVTGSSDNEYFVVDPREY 599
QY      600 DLKWEFPRENLEFGKVLGSGAFGKVMNATAYGISTGVSIQVAVKMLKEKADSSREALM 659
Db      600 DLKWEFPRENLEFGKVLGSGAFGKVMNATAYGISTGVSIQVAVKMLKEKADSSREALM 659
QY      660 SELKMTQLGSHENIVNLLGACTLSGPIYLIFFECCYGDLLNLYRSKREKPHRTWTEIFK 719
Db      660 SELKMTQLGSHENIVNLLGACTLSGPIYLIFFECCYGDLLNLYRSKREKPHRTWTEIFK 719
QY      720 EHNFSFYPTQSHPNSSMPGSRVQIHPDSDQISGLHNSPHSDEIEYENOKREED 779
Db      720 EHNFSFYPTQSHPNSSMPGSRVQIHPDSDQISGLHNSPHSDEIEYENOKREED 779
QY      780 LNVLTPEDLFCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSD 839
Db      780 LNVLTPEDLFCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSD 839
QY      840 SNYVVRGNARLPVKWAPESLFEIGYTIKSDVMSYGILLWEIFSLGVNPPYGPVDANFY 899
Db      840 SNYVVRGNARLPVKWAPESLFEIGYTIKSDVMSYGILLWEIFSLGVNPPYGPVDANFY 899
QY      900 KLIQNGFMDQPPFATEEIIYIMQSCWAFDSKRKPSFNLTSFLGCOLADAEAMQVND 959
Db      900 KLIQNGFMDQPPFATEEIIYIMQSCWAFDSKRKPSFNLTSFLGCOLADAEAMQVND 959
QY      960 GRVSECPHTYQNRPPFSREMDLGLLSPQAQVEDS 993
Db      960 GRVSECPHTYQNRPPFSREMDLGLLSPQAQVEDS 993

RESULT 13
US-08-222-299-2
; Sequence 2, Application US/08222299
; Patent No. 5635388
; GENERAL INFORMATION:
; APPLICANT: Bennett, Brian D.
; APPLICANT: Broz, Susan D.
; APPLICANT: Matthews, William
; APPLICANT: Zeigler, Francis C.
; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES TH
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080

```

COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/222,299  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Hasak, Janet E.  
REGISTRATION NUMBER: 28,616  
REFERENCE/DOCKET NUMBER: 879  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1896  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1000 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

US-08-222-299-2

Query Match 86.0%; Score 4533.5; DB 1; Length 1000;  
Best Local Similarity 85.6%; Pred. No. 0;  
Matches 851; Conservative 56; Mismatches 84; Indels 3; Gaps 2;

QY 1 MPALA-RDAGTVLLVVFSAIFGTITNQLPVKVLINHKNDSSVGSXSSYPVMS 59  
DB 1 MRALAQRDRRLVLLVLSWILETVNQLPVKVLINSHENNGSAGKPSYRVRGS 60

QY 60 PEDLGCALRQSSGTVYAAAVEVDVSASITLQVLVDAPGNI SCLWYFKHSSLCQPHD 119  
DB 61 PEDLQCTPRQSGTVYEAATVEAESGITLQVLATPGDLSCLWYFKHSSLCQPHD 120

QY 120 LQNRGVSVMLKMTQAGEYLLFTOSEATNTYLTFTVSRINTLLTLRPPYRKMENQ 179  
DB 121 LQNRGVSMALNVTTQAGEYLLHIOSEAAVTYLTFTVNRDPTQLVLRPPYRKMENQ 180

QY 180 DALVCISEVPDIVEHVLCDQSGESCKERSPAVVKKEKVLHFLGTDIRCCARNLGR 239  
DB 181 DALLCISEGPEPTVEHVLCDSSHRESCKERSPAVVKKEKVLHFLGTDIRCCARNLGR 240

QY 240 ECTRLFTIDLNQPTTLPOLFLKVGEPMLRCKAVHVNHGFLTWELNKALBEGNYFE 299  
DB 241 ESTKLFTIDLNQAPSTLPOLFLKVGEPMLRCKAIVHVNHGFLTWELDKALEGSYFE 300

QY 300 MSTYSTNRTMIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVKGFNATNSSEY 359  
DB 301 MSTYSTNRTMIRILFAFVSSVGRNDTGYTCSSSKHPSQSALVTILEKGFNATNSSEY 360

QY 360 EIDYEFCSVRKAPQIRCTWTFRSKSPCEQKGLDNGYSIKFCNKHGQGEVIEH 419  
DB 361 EIDPYEKFCSVRKAPQIRCTWTFRSKSPCEQKGLDNGYSISKFDHKNKQGEVIFY 420

QY 420 AENDDAQFTKMTLIRKPKOVLAESAQSCFSDGYPLPSWTWKCKSDKSPNCTBEIT 479  
DB 421 AENDDAQFTKMTLIRKPKOVLANASQSCSDGYPLPSWTWKCKSDKSPNCTBEIP 480

QY 480 EGVNKRANRKFQGWSSSTLNSEAIKGLVRCCKAYNSLGTSCETILLNSPGFPPIQ 539  
DB 481 EGVNKRANRKFQGWSSSTLNSEAGKGLLVKCCAYNSMGTSCTETILLNSPGFPPIQ 540

QY 540 DNISFYATIGVCLLFIWVLLIHKYKQFRYSQLOQWQVGTSSDNEYFYVDFREY 599  
DB 541 DNISFYATIGVCLLFIWVLLIHKYKQFRYSQLOQWQVGTSSDNEYFYVDFREY 600

QY 600 DLKWEFFRENLEFGKVLGSGAFGRVNMATAYGISTGVSIQVAVKMLKEKADSSEREALM 659

DB 601 DLKWEFFRENLEFGKVLGSGAFGRVNMATAYGISTGVSIQVAVKMLKEKADSSEREALM 660  
QY 660 SELKMTQLGSHENIVNLLGACTLSQPIYLIPEYCCYDGLLNVLRSKREKFTWTEIFK 719  
DB 661 SELKMTHLGHNDIVNLLGACTLSGPVYLIPEYCCYDGLLNVLRSKREKFTWTEIFK 720  
QY 720 EHNFSFYPTFQSHPNSSMPGSRVQIHPDQISGLHNSFHSEDEIEYENQKRL--EEE 777  
DB 721 EHNFSFYPTFQSHPNSSMPGSRVQIHPDQISGLHNSFHSEDEIEYENQKRLAEEEE 780  
QY 778 EDNLVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGVVVKICDFGLARDIM 837  
DB 781 EDNLVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGVVVKICDFGLARDIL 840  
QY 838 SDSNYVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNPPYGPIDAN 897  
DB 841 SDSNYVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNPPYGPIDAN 900

QY 898 FYKLIONGKMDOPFYATEIYIIMOSWAFDSRKPSPFNLTSLFCOLADAEAMYQN 957  
DB 901 FYKLIOGFRMEQPFYATEGIYFVMSWAFDSRKPSPFNLTSLFCOLADAEAMYQN 960

QY 958 VDGVRSECHTYQNRPPFSREMDLGLLSPQAOVE 991  
DB 961 MGNVPEHPSIYQNRRLPLSREAGSEPPSPQAOVK 994

RESULT 14

US-08-434-878-2  
Sequence 2, Application US/08434878  
Patent No. 5997865  
GENERAL INFORMATION:  
APPLICANT: Bennett, Brian D.  
APPLICANT: Broz, Susan D.  
APPLICANT: Matthews, William  
APPLICANT: Zeigler, Francis C.  
TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES TH  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/434,878  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Hasak, Janet E.  
REGISTRATION NUMBER: 28,616  
REFERENCE/DOCKET NUMBER: 879  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1896  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1000 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-434-878-2

Query Match 86.0%; Score 4533.5; DB 2; Length 1000;

Best Local Similarity 85.6%: Pred. No. 0;		Matches 851; Conservative 56; Mismatches 84; Indels 3; Gaps 2	
Qy	1	MPALA-RDAGTVPLLVAFESAMIFGTITNQDLPIVKCVILINHKNDSSVGKSSSPYMWSES	59
Db	1	MRLAQRSDRELLLLVVLVSMLEIVTNQDLPIVKCVILISHENNGSSAGKSPSSYMYRGS	60
Qy	60	PEDLGCLARPOSGGTVEYAAAVVDVDSASITLQVLVDAPGNISCLVWPKHSGSLMCQPHFD	119
Db	61	PEDLQCTPRROSEGTVYEAATVEAESGSITLQVLATPGDLLSCLVWPKHSGSLGCQPHFD	120
Qy	120	LONRGVSMVTLKMTETQAGYLLFIQSEANTYITLFTVSRNTLLYTLRPPYFRKEMNQ	179
Db	121	LONRGVSMALNVTETQAGYLLHIQSEANYTVLFTVNRDQLYVLRPPYFRKEMNQ	180
Qy	180	DALVCISSEVPEPIVWVLCDSQGESCKEESPAVVYKKEEKVLHELFGTDIRCCARNELGR	239
Db	181	DALLCISEGVPEPTVWVLCSSHRESCKEESPAVVYKKEEKVLHELFGTDIRCCARNALGR	240
Qy	240	ECTBLFTIDLNTQTTLPOLFLVGEPLWTRCKAVHVNHGFLTWELNKALEBGNVFE	299
Db	241	ESTKLFTIDLNAQPSLTPOLFLVGEPLWTRCKAIHVNHGFLTWELNKALEBGSYFE	300
Qy	300	MSTYSTNRTMIRILFAPVSSVARNDGTYYTCCSKHPSQSALVTIVKGFTINATNSSDY	359
Db	301	MSTYSTNRTMIRILFAPVSSVGRNDGTYYTCCSKHPSQSALVTILEKGFINATSSQEY	360
Qy	360	EIDQYEFCSVRPKAYPIQIRCTWTFRSKSPCEQGLDNGYSISKFNHKKHQGEYIFH	419
Db	361	EIDPYEXFCSVRPKAYPIRCTWIFSQAQSPCEQGLEDDGYSISKFCDHNKPKGEYIFY	420
Qy	420	AENDDAQFTKMTNIRKPOVLAEASQASQSCSDGYPPLPSWTWKKCSKSPNCTEIT	479
Db	421	AENDDAQFTKMTNIRKPOVLANASQASQSCSDGYPPLPSWTWKKCSKSPNCTEIP	480
Qy	480	EGVWNRKANRVFGQWSSSTLNLWSEATKGLVKCCAVNSLGTSCETILLNSPGPFPIQ	539
Db	481	EGVWNKANRVFGQWSSSTLNLWSEAGLLVKCCAINSGTSCETILLNSPGPFPIQ	540
Qy	540	DNISFYATIGVCLLFIIVLTLLICHYKKQPRYESQLQMVQVGTSSDNEYFYFVDFREY	599
Db	541	DNISFYATIGLCLFPIVLVILICHYKKQPRYESQLQMIQVTPGLDNEYFYVDFRDY	600
Qy	600	DLKWEFPRENLEFGKVLGSGAFGKVMNATAYIGISKTGVSIOAVKMLKEKADSSEREALM	659
Db	601	DLKWEFPRENLEFGKVLGSGAFGRVMNATAYIGISKTGVSIOAVKMLKEKADSCKEALM	660
Qy	660	SELKQMTQLGSHENIVNLGACTLSGPYLLIPEYCCYGDLNLYRSKREKPHRTWTETFK	719
Db	661	SELKQMTLGHHDNIVNLGACTLSGPYLLIPEYCCYGDLNLYRSKREKPHRTWTETFK	720
Qy	720	EHNFSFYPTFQSHPSNMPGSRREYQIHPDSDQISGLHGNSPHSDEIEYENQKEL--EEE	777
Db	721	EHNFSFYPTQAHNSNMPGSRREYQLHPDLDQSGFNGNSIHSDEIEYENQKLAEEEE	780
Qy	778	EDLNVLTPEDLLCFAYQVAKGMEFLPKSCVHRDLAARNVLVTHGKVVKICDFGLARDIM	837
Db	781	EDLNVLTFEDLLCFAYQVAKGMEFLPKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL	840
Qy	838	SDSNVYVRGNARLPEVKMAPELSEGIYTIKSDVMSYGILLWEIFSLGVNPPGIPVDAN	897
Db	841	SDSNVYVRGNARLPEVKMAPELSEGIYTIKSDVMSYGILLWEIFSLGVNPPGIPVDAN	900
Qy	898	FYKLIQNGFKMDQPFYATEEYIIMQSCWAFDSRKRPSFNNLTSLGCOLADASEAMYNQ	957
Db	901	FYKLIQSGFKMEQPFYATEGIYFWQSCWAFDSKRPSFNNLTSLGCOLAEEAEMYNQ	960
Qy	958	VDGRVSECPHTYQNRPRFSRMDLGLSLSPQAQVE	991
Db	961	MGNVNPBPHPSIYQNRRLPSREAGSEPPSPQAQVK	994

RESULT 15

PCT-US95-03718-2  
; Sequence 2, Application PC/TUS9503718  
; GENERAL INFORMATION:  
; APPLICANT: GENENTECH, INC.  
; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AN  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/03718  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Wendy M. Lee  
; REGISTRATION NUMBER: 00,000  
; REFERENCE/DOCKET NUMBER: 879PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-1994  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1000 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
PCT-US95-03718-2

Query Match	86.0%;	Score 4533.5;	DB 5;	Length 1000;
Best Local Similarity	85.6%;	Pred. No. 0;		
Matches 851;	Conservative 56;	Mismatches 84;	Indels 3;	Gaps 2
QY	1	MPALA-RDAGTVPLLVVFSAMIFGTITNQDLPVTKCVLIINHKNDDSSVKGSSSPWMS	59	
DB	1	MRALAQRSDRLILLVLSWILETTVTDLPVTKCVLIISHENNGSSACKPSSYRNVRGS	60	
QY	60	PEDGCGALRPQSSGTVYEAFAVEVDVSASITLQVLVDAPGNISCLWVPKHSNLNCQPHFD	119	
DB	61	PEDLQCTPRQSGTVYEAATVEAESSITLQVOLATPGDLSCLWVPKHSNLGQCPHFD	120	
QY	120	LQNRGVSVVLAKMTETOAGEYILLFTQSEATNYTILFTVSTFNTLLYTLRRPYPKEMNQ	179	
DB	121	LQNRGVSMALLNVTETOAGEYILLHIQSEAAAYTLVFTVNVDRDQLVYLRPPYFRMENO	180	
QY	180	DALVCISESPYPIVEVLVDCSQSCKESBPVAVKKEKVLHFGTDIRCCARNEILGR	239	
DB	181	DALLCISEGPEPTVEWLVCSSHRESCKESGPAVVRKEEKLHFGTDIRCCARNALGR	240	
QY	240	ECTRLFTIDLNOTPTTLPOLFLKVGGEPLWTRCKAVHVNHGFGLTWELNKALEBGNVFE	299	
DB	241	ESTKLFTIDLNQAPQSTLPOLFLKVGGEPLWIRCKAIHVNHGFGLTWELDKALEBGSYFE	300	
QY	300	MSYSTNRTMIRILLAFVSSVARNDTGYYTCSSSKHPSQSALVTIVGKGFINATNSSDDY	359	
DB	301	MSYSTNRTMIRILLAFVSSVGRNDTGYYTCSSSKHPSQSALVTILEKGFINATSSQEEY	360	
QY	360	EIDQYEEFCFSVRFKAYPQIRCTWTWTSRKSFPCEBKGLDNGYSISKFCNHKGHPGEYIFH	419	
DB	361	EIDPYBKPCFSVRFKAYPRIRCTWTWISQAGFPCEBQRGLEDGYSISKFCDHKNKPGEYIFY	420	
QY	420	AENDDAQFTKMTFINIRRRKQVLAESAASQASCFSDGYPILPSMTWTKGSKDSKSPNCTEIBT	479	

Db	421	AEHDDAQTFTMFTLNIRKPPQVLANASASQSCSDGYPLPSWTWKKCDKSPNCTEIP	480
Qy	480	EGVNRKANRKFQOWSSSTLNSEAIKGFVKCCAYNSLGTSCETILLNSPGPPPIQ	539
Db	481	EGVNRKANRKFQOWSSSTLNSEAGLGLVKCCAYNSMGTSCTIFLNSPGPPPIQ	540
Qy	540	DNISFYATIGVCLLFIWVLTLLIHKYKQPRYSOLOMVOVTCSSNEYFYVDFREY	599
Db	541	DNISFYATIGLCLFPIVVLIVLIHKYKQPRYSOLOMVOVTCSSNEYFYVDFREY	600
Qy	600	DLKWEPPRENLEFGKVLGSGAFGVMATAGISKTGVSIOVAVKMLKEKADSSERBALM	659
Db	601	DLKWEPPRENLEFGKVLGSGAFGVMATAGISKTGVSIOVAVKMLKEKADSSERBALM	660
Qy	660	SELKMMTQLGSHENIVNLLGACTLSGPIYLIFEXCYGDLNLYLRSKREKPHRTWTETFK	719
Db	661	SELKMMTHLGHHDNIVNLLGACTLSGPIYLIFEXCYGDLNLYLRSKREKPHRTWTETFK	720
Qy	720	EHNFSPYPTQSHNSNMPGSRVQIHPDSDQISGLHGNFSHSEDEIYENQKRL--EE	777
Db	721	EHNFSPYPTQAHNSNMPGSRVQIHPDSDQISGLHGNFSHSEDEIYENQKRLAEEEE	780
Qy	778	EDLNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIM	837
Db	781	EDLNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL	840
Qy	838	SDSNVVRGNARLPVKWMAPELSEFEGYITIKSDVMSYGILLWEIFSLGVNFPYGPVVDAN	897
Db	841	SDSNVVRGNARLPVKWMAPELSEFEGYITIKSDVMSYGILLWEIFSLGVNFPYGPVVDAN	900
Qy	898	FYKLIQSGFKMDQPFYATEEYIIMQSCWAFDSRKRPSFPNLTSLGCOLADAEAMYQN	957
Db	901	FYKLIQSGFKMDQPFYATEEYIIMQSCWAFDSRKRPSFPNLTSLGCOLADAEAMYQN	960
Qy	958	VDGRVSECPHTYQNRPRPSREMDLGLLSPOAOVE	991
Db	961	MGGNVPEHPSIYQNRRLPSREAGSEPPSPQAOVK	994

Search completed: August 26, 2003, 07:29:55  
 Job time : 25.5108 secs